

---

<i>Address (NTU)</i>	Block N4, 02c-79 <a href="#">NTU-NLP Group</a> <a href="#">School of Computer Science and Engineering (SCSE)</a> <a href="#">Nanyang Technological University (NTU)</a>
<i>Salesforce</i>	Floor 40, Suntec City Tower 2 <a href="#">Salesforce AI Research</a>
<i>Phone</i>	(+65) 90696196, 69041107
<i>Mail</i>	<a href="mailto:srjoty@ntu.edu.sg">srjoty@ntu.edu.sg</a> , <a href="mailto:sjoty@salesforce.com">sjoty@salesforce.com</a>
<i>WWW</i>	<a href="https://raihanjoty.github.io/">https://raihanjoty.github.io/</a>
<i>Nationality</i>	<b>Canadian</b>
<i>h-index</i>	<b>35</b> (Google Scholar)

---

## RESEARCH INTERESTS

- Natural Language Processing
  - NLP tools (Syntactic/Semantic/Discourse Parsing, Coherence models, Topic models)
  - Multi-lingual NLP (Machine Translation, Cross-lingual tasks)
  - Text generation (text2text, data2text)
  - Robust and fair NLP
  - NLP Applications (Question answering, Dialogue systems)
  - Multi-modal NLP (Visual-Language Pretraining, Speech, NLP for programming)
- Machine Learning & Data Science
  - Deep/Representation learning
  - Probabilistic graphical models
  - Reinforcement learning

---

## PROFESSIONAL EMPLOYMENTS

<b>Assistant Professor</b> <i>Nanyang Technological University (NTU), Singapore</i> <i>Founder of NTU-NLP Group</i>	2017-present
<b>Senior Manager</b> <i>Salesforce AI Research, Singapore</i> <i>Founding member of the Asia lab</i>	2019-present
<b>Research Scientist</b> <i>Qatar Computing Research Institute, Doha, Qatar</i>	2014-2017

<b>Research &amp; Teaching Assistant</b> <i>University of British Columbia</i> <i>University of Lethbridge</i>	2006-2013
<b>Research Intern</b> <i>Microsoft Research</i>	2010
<b>Lecturer</b> <i>Islamic University of Technology, Dhaka, Bangladesh</i>	2005-2006

## EDUCATION

---

<b>Ph.D. (Computer science)</b> <i>University of British Columbia (UBC), Vancouver, Canada</i> “Discourse analysis of asynchronous conversations” Thesis advisors: <a href="#">Giuseppe Carenini</a> & <a href="#">Raymond T. Ng</a> Internal committee members: <a href="#">Nando de Freitas</a> (now at Oxford/DeepMind) & <a href="#">Rachel Pottinger</a> External committee members: <a href="#">Carolyn P. Rosé</a> (CMU) & <a href="#">Amanda Stent</a> (Yahoo! Research)	2008-2013
<b>M.Sc. (Computer science)</b> <i>University of Lethbridge (UofL), Alberta, Canada</i> “Answer Extraction for Simple and Complex Questions” Thesis advisor: <a href="#">Yllias Chali</a> Other committee members: Sajjad Zahir, Howard Cheng & Giuseppe Carenini (external)	2006-2008
<b>B.Sc. (Computer science)</b> <i>Islamic University of Technology (IUT), Dhaka, Bangladesh</i> CGPA: 4.97/5.00 (2nd in the class)	2001-2005

## MAJOR HONORS AND AWARDS

---

- [NSERC Alexander Graham Bell Canada Graduate Scholarship<sup>1</sup>](#) (CGS-D)  
2010 – 2012                      \$35,000 CAD/Year
- [Four Year Fellowship \(UBC\)](#)  
2010 – 2013                      \$18,000 CAD/Year
- [Faculty of Science Graduate Studies Award \(UBC\)](#)  
2010 – 2013                      \$7,200 CAD/Year
- [Nominated for MIT Technology Review’s Innovators Under 35 \(Asia\)](#) 2017 – 2018
- Microsoft Research Excellent Intern, 2010
- Nomination for Governor General Gold Medal for Best MSc. Thesis (UofL), 2008
- David R. Cheriton Scholarship, University of Waterloo, 2008 [Declined]

---

<sup>1</sup>This award is given only to the top-ranked PhD students across Canada.

- FS Chia Ph.D. Scholarship, University of Alberta, 2008 [Declined]
- First class 2nd in BSc, 5th in higher secondary (HSC) & 11th in secondary (SSC) merit lists

## TEACHING EXPERIENCE

---

**Instructor, Nanyang Technological University (NTU)**

2017-

- **Year 2021 - 2022:**
  - (i) Term 1: CZ4045: Natural Language Processing (Lec, Tut); CZ2002: Object Oriented Design & Programming (Tut. & Lab)
  - (ii) Term 2: CE-7455 (PhD level): *Deep Learning for NLP* (added meta learning and VAEs)
- **Year 2020 - 2021:**
  - (i) Term 1: CZ4045: Natural Language Processing (*syllabus revised with deep learning*) (Lec, Tut); CZ2002: Object Oriented Design & Programming (Tut. & Lab)
  - (ii) Term 2: CE-7455 (PhD level): *Deep Learning for NLP* (added multilingual NLP)
- **Year 2019 - 2020:**
  - (i) Term 1: CZ3005: Artificial Intelligence (Lec, Lab, Tut); CZ1007: Data Structures & Algorithm (Tut. & Lab)
  - (ii) Term 2: CE-7455 (PhD level): *Deep Learning for NLP* (*new offer*)
- **Year 2018 - 2019:**
  - (i) Term 1: CZ3005: Artificial Intelligence (Lec, Lab, Tut), CZ1007: Data Structures & Algorithm (Tut. & Lab)
  - (ii) Term 2: CZ3005: Artificial Intelligence (Lec, Lab, Tut)
- **Year 2017 - 2018:**
  - (i) Term 1: CZ1007: Data Structures & Algorithm (Tutorial & Lab)
  - (ii) Term 2: CZ3005: Artificial Intelligence (Lecture, Lab, & Tutorial)

**Teaching Assistant, University of British Columbia**

2008-2013

- Courses TAed: Grad NLP, Grad/Ugrad AI, Object Oriented Programming, Intro. to Comp.

**Teaching Assistant, University of Lethbridge**

2006-2008

- Courses TAed: Introduction to Computers.

**Lecturer, Islamic University of Technology**

2005-2006

- Courses Taught: Theory of computation, UNIX prog., Systems prog., OOP.

## PATENTS

---

### Granted Patents

1. Xuan-Phi Nguyen, **Shafiq Joty**, and Chu Hong Hoi., “[SYSTEMS AND METHODS FOR A TRANSFORMER NETWORK WITH TREE-BASED ATTENTION FOR NATURAL LANGUAGE PROCESSING](#)”, **US 2021/0049236 A1**, US Patent (2021).
2. Karan Aggarwal, Swaraj Khadanga, **Shafiq Joty**, Louis Kazaglis, and Jaideep Srivastava, “[COMPUTER ARCHITECTURE FOR IDENTIFYING SLEEP STAGES](#)”, **US 2021/0023331 A1**, US Patent (2021).
3. Karan Aggarwal, Swaraj Khadanga, **Shafiq Joty**, and Jaideep Srivastava, “[USING CLINICAL NOTES WITH TIME SERIES DATA FOR ICU MANAGEMENT](#)”, *NTU REF: 2019-374* **US Patent**, (2020).
4. Samson Tan and **Shafiq Joty.**, “[SYSTEMS AND METHODS FOR GENERATING NATURAL LANGUAGE PROCESSING TRAINING SAMPLES WITH INFLECTIONAL PERTURBATIONS](#)”, **US 2021/0173872 A1**, US Patent (2021).
5. Yifan Gao, Chu Hong Hoi, **Shafiq Joty** and Chien-Sheng Wu., “[SYSTEMS AND METHODS FOR EXPLICIT MEMORY TRACKER WITH COARSE-TO-FINE REASONING IN CONVERSATIONAL MACHINE READING](#)”, *Patent Nb: 16/863,999* **US2021/0174023 A1**, US Patent (2021).
6. Yue Wang, Chu Hong Hoi, and **Shafiq Joty.**, “[UNIFIED VISION AND DIALOGUE TRANSFORMER WITH BERT](#)”, *Patent Nb: 16/929,738* **US2021/0232773 A1**, US Patent (2021).
7. Weishi Wang, **Shafiq Joty**, and Chu Hong Hoi., “[SYSTEMS AND METHODS FOR RESPONSE SELECTION IN MULTI-PARTY CONVERSATIONS WITH DYNAMIC TOPIC TRACKING](#)”, *Patent Nb: 17/014,458* **US Patent**, Filing Date: September 8, 2020 (2020).

### Filed Patents

1. Amrita Saha, **Shafiq Joty**, and Chu Hong Hoi., “SYSTEMS AND METHODS FOR NUMERICAL REASONING BY A PARTIALLY SUPERVISED NUMERIC REASONING MODULE NETWORK ”, *Patent Nb: 17/162,289* **US Patent**, Filing Date: 01/29/2021 (2021).
2. Samson Tan and **Shafiq Joty.**, “SYSTEMS AND METHODS FOR CODE-MIXING ADVERSARIAL TRAINING”, *Patent Nb:17/150,988* **US Patent**, Filing Date: January 15, 2021 (2021).
3. Yue Wang, Weishi Wang, **Shafiq Joty** and Chu Hong Hoi., “SYSTEMS AND METHODS FOR CODE UNDERSTANDING AND GENERATION”, *70689.170PV01 (63/189,857)* **US patent**, Filing Date: May 18, 2021 (2021).
4. Akhilesh Deepak Gotmare, Junnan Li, **Shafiq Joty** and Chu Hong Hoi., “SYSTEMS AND METHODS FOR SEMANTIC CODE SEARCH”, *70689.166PV01 (63/189,854)* **US patent**, Filing Date: May 18, 2021 (2021).
5. DING Bosheng, LIU Linlin, BING Lidong, NGUYEN Hai Thien, **Joty Shafiq**, MIAO Chunyan, Si LUO., “Method And Apparatus For Training A Text Tagging Model”, *202011075457.0. (NTU Ref: 2021-023)* **Chinese patent**, Filing Date: 9 October (2021).

6. Yingzhu Zhao\*, Chongjia Ni, Cheung-Chi Leung, **Shafiq Joty**, Eng Siong, and Bin Ma, “Speech Recognition System, Method, Apparatus And Device.”, *NTU Ref: 2021-077-01-CN Chinese patent*, Filing Date: 22 February (2021).

## IMPACT AND VISIBILITY

---

\*\* Most of the codes of my papers can be found in the [NTU-NLP](#) group page

## PEER-REVIEWED JOURNAL PAPERS

---

\* indicates co-lead author

1. Tasnim Mohiuddin and **Shafiq Joty**, “[Unsupervised Word Translation with Adversarial Autoencoder](#)”, *Computational Linguistics* **46:2**, (presented at ACL-2020) (Jun, 2020).
2. Xiangxi Shi, Jianfei Cai, Jiuxiang Gu, and **Shafiq Joty**., “[Video Captioning with Boundary-Aware Hierarchical Language Decoding and Joint Video Prediction](#)”, *Neurocomputing* **417**, 347-356 (Dec, 2020).
3. Jing Ma, Wei Gao, **Shafiq Joty**, and Kam-Fai Wong, “[An Attention-based Rumor Detection Model with Tree-structured Recursive Neural Networks](#)”, *ACM Transactions on Intelligent Systems and Technology (TIST)* **11:04**, (42) (2020).
4. Lorainne Car, Dhakshenya Dhinakaran, Bhone Kyaw, Tobias Kowatsch, **Shafiq Joty**, Yin Theng, and Rifat Atun, “[Conversational agents in healthcare: a scoping review and conceptual analysis](#)”, *Journal of Medical Internet Research (JMIR)* **22:8**, (Aug, 2020).
5. **Shafiq Joty** and Tasnim Mohiuddin, “[Modeling Speech Acts in Asynchronous Conversations: A Neural-CRF Approach.](#)”, *Computational Linguistics (Special Issue on Language in Social Media)* **44:4**, 859 - 894 (2018).
6. **Shafiq Joty**, Francisco Guzmán, Lluís Màrquez, and Preslav Nakov., “[Discourse Structure in Machine Translation Evaluation.](#)”, *Computational Linguistics.* **43:4**, 683–722 (Dec, 2017).
7. **Shafiq Joty**, Nadir Durrani, Hassan Sajjad, and Ahmed Abdelali., “[Domain Adaptation Using Neural Network Joint Model.](#)”, *Journal of Computer Speech and Language (Special Issue on Deep Learning for Machine Translation).* **45:C**; **csl.2016.12.006**, 161–179 (Sep, 2017).
8. Francisco Guzmán, **Shafiq Joty**, Lluís Màrquez, and Preslav Nakov., “[Machine Translation Evaluation with Neural Networks.](#)”, *Journal of Computer Speech and Language (Special Issue on Deep Learning for Machine Translation).* **45:C**; **csl.2016.12.005**, 180–200 (Sep, 2017).
9. Aarti Sathyanarayana, **Shafiq Joty**, Luis Fernandez-Luque, Ferda Ofli, Jaideep Srivastava, Ahmed Elmagarmid, Shahrads Taheri, and Teresa Arora., “[Sleep Quality Prediction From Wearable Data Using Deep Learning](#)”, *JMIR Mhealth Uhealth (JMU)* **(4(4):e125)**, JMIR (2016).
10. **Shafiq Joty**, Giuseppe Carenini and Raymond Ng, “[CODRA: A Novel Discriminative Framework for Rhetorical Analysis.](#)”, *Computational Linguistics.* **41:3**, 385–435, MIT press (2015).

11. **Shafiq Joty**, Giuseppe Carenini and Raymond Ng, “[Topic Segmentation and Labeling in Asynchronous Conversations.](#)”, *Journal of AI Research*. **47**, 521-573 (2013).
12. Yllias Chali\*, Sadid Hasan\* and **Shafiq Joty\***, “[Improving graph-based random walks for complex question answering using syntactic, shallow semantic and extended string subsequence kernels.](#)”, *Information Processing and Management*. **47:6**, 843–855, Elsevier (2011).
13. Yllias Chali\*, **Shafiq Joty\*** and Sadid Hasan, “[Complex Question Answering: Unsupervised Learning Approaches and Experiments.](#)”, *Journal of AI Research*. **35**, 1-47 (2009). (I was a co-lead author of this paper.)

## CONFERENCE PAPERS UNDER REVIEW

---

## PEER-REVIEWED CONFERENCE PAPERS [BY YEAR]

---

\* indicates co-lead author

### 2022

1. Chengwei Qin and **Shafiq Joty**, “[LFPT5: A Unified Framework for Lifelong Few-shot Language Learning Based on Prompt Tuning of T5](#)”, In *International Conference on Learning Representations (ICLR-2022)*, (Online).
2. Xuan-Phi Nguyen, Hongyu Gong, Yun Tang, Changhan Wang, Philipp Koehn and **Shafiq Joty**, “[Contrastive Clustering to Mine Pseudo Parallel Data for Unsupervised Translation](#)”, In *International Conference on Learning Representations (ICLR-2022)*, (Online).
3. Amrita Saha, **Shafiq Joty**, and Steven Hoi., “[Weakly Supervised Neuro-Symbolic Module Networks for Numerical Reasoning.](#)”, In *Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-2022)*, Vancouver, Canada (2022).
4. Jiahui Gao, Yi Zhou, Philip Yu, **Shafiq Joty**, and Jiuxiang Gu., “[Unsupervised Cross-lingual Image Captioning.](#)”, In *Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-2022)*, Vancouver, Canada (2022).

### 2021

1. Junnan Li, Ramprasaath R., Akhilesh Deepak, **Shafiq Joty**, Caiming Xiong, and Steven Hoi., “[Align before Fuse: Vision and Language Representation Learning with Momentum Distillation.](#)”, In *2020 Conference on Neural Information Processing Systems (NeurIPS’21 (Spotlight ~3%))*, Online (2021).
2. Xiang Lin, Simeng Han and **Shafiq Joty.**, “[Straight to the Gradient: Learning to Use Novel Tokens for Neural Text Generation.](#)”, In *Proceedings of the Thirty-eighth International Conference on Machine Learning (ICML’21 (Long talk ~3%))*, Virtual Conference (2021).
3. Xuan-Phi Nguyen, **Shafiq Joty**, Wu Kui and Ai Ti Aw., “[Cross-model Back-translated Distillation for Unsupervised Machine Translation.](#)”, In *Proceedings of the Thirty-eighth International Conference on Machine Learning (ICML’21)*, Virtual Conference (2021).

4. M Saiful Bari and Tasnim Mohiuddin and **Shafiq Joty**, “[UXLA: A Robust Unsupervised Data Augmentation Framework for Cross-Lingual NLP.](#)”, *In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL’21)*, Bangkok, Thailand (2021).
5. Samson Tan, **Shafiq Joty**, Kathy Baxter, Araz Taeihagh, Gregory A. Bennett and Min-Yen Kan, “[Reliability Testing for Natural Language Processing Systems: An Adversarial Perspective.](#)”, *In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL’21)*, Bangkok, Thailand (2021).
6. Thanh-Tung Nguyen, Xuan-Phi Nguyen, **Shafiq Joty** and Xiaoli Li., “[A Conditional Splitting Framework for Efficient Constituency Parsing.](#)”, *In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL’21)*, Bangkok, Thailand (2021).
7. Linlin Liu, Bosheng Ding, Lidong Bing, **Shafiq Joty**, Luo Si and Chunyan Miao, “[MulDA: A Multilingual Data Augmentation Framework for Low-Resource Cross-Lingual NER.](#)”, *In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL’21)*, Bangkok, Thailand (2021).
8. Tasnim Mohiuddin, M Saiful Bari and **Shafiq Joty.**, “[AugVic: Exploiting BiText Vicinity for Low-Resource NMT.](#)”, *In Findings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL’21)*, Bangkok, Thailand (2021).
9. Yue Wang, Weishi Wang, **Shafiq Joty**, and Steven Hoi., “[CodeT5: Identifier-aware Unified Pre-trained Encoder-Decoder Models for Code Understanding and Generation.](#)”, *In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21)*, Punta Cana, Dominican Republic (2021).
10. Tao Yu and **Shafiq Joty**, “Effective Fine-tuning Methods for Cross-lingual Adaptation.”, *In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21)*, Punta Cana, Dominican Republic (2021).
11. Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, **Shafiq Joty**, Eng Siong, and Bin Ma., “A Unified Speaker Adaptation Approach for ASR.”, *In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21)*, Punta Cana, Dominican Republic (2021).
12. Ben Krause, Akhilesh Deepak, Bryan McCann, Nitish Shirish, **Shafiq Joty**, Richard Socher, and Nazneen Fatema., “[GeDi: Generative Discriminator Guided Sequence Generation.](#)”, *In Findings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21)*, Punta Cana, Dominican Republic (2021).
13. Samson Tan and **Shafiq Joty**, “[Code-Mixing on Sesame Street: Dawn of the Adversarial Polyglots.](#)”, *In Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-2021)*, Mexico City, Mexico (2021).
14. Thanh-Tung Nguyen, Xuan-Phi Nguyen, **Shafiq Joty**, and Xiaoli Li, “[RST Parsing from Scratch.](#)”, *In Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-2021)*, Mexico City, Mexico (2021).
15. Alexander Fabbri, Simeng Han, Haoyuan Li, Haoran Li, Marjan Ghazvininejad, **Shafiq Joty**, Dragomir Radev, and Yashar Mehdad, “[Improving Zero and Few-Shot Abstractive](#)



- Summarization with Intermediate Fine-tuning and Data Augmentation.”, *In Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-2021)*, Mexico City, Mexico (2021).
16. Weiwen Xu, AiTi Aw, Yang Ding, Kui Wu and **Shafiq Joty**, “[Addressing the Vulnerability of NMT in Input Perturbations.](#)”, *In Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Industry track) (NAACL-2021)*, Mexico City, Mexico (2021).
  17. Tasnim Mohiuddin\*, Prathyusha Jwalapuram\*, Xiang Lin\*, and **Shafiq Joty\***, “[Rethinking Coherence Modeling: Synthetic vs. Downstream Tasks.](#)”, *In Proceedings of European Chapter of the Association for Computational Linguistics (EACL-2021)*, Kyiv (2021).
  18. Xiangju Li, Wei Gao, Shi Feng, Wang Daling, and **Shafiq Joty**, “[Span-Level Emotion Cause Analysis by BERT-based Graph Attention Network.](#)”, *In Proceedings of The 30th ACM International Conference on Information and Knowledge Management (CIKM-2021)*, Online (2021).
  19. Xiangju Li, Wei Gao, Shi Feng, Wang Daling, and **Shafiq Joty**, “[Span-level Emotion Cause Analysis with Neural Sequence Tagging.](#)”, *In Proceedings of The 30th ACM International Conference on Information and Knowledge Management (CIKM-2021)*, Online (2021).
  20. Weiguo Wang, Sourav S Bhowmick, Hui Li, **Shafiq Joty** and Siyuan Liu., “[Towards Enhancing Database Education: Natural Language Generation Meets Query Execution Plans.](#)”, *In Proceedings of 2021 ACM SIGMOD International Conference on Management of Data (SIGMOD-2021)*, Shaanxi, China (2021).
  21. Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, **Shafiq Joty**, Eng Siong, and Bin Ma., “[Preventing Early Endpointing for Online Automatic Speech Recognition.](#)”, *In International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2021)*, Shanghai, China (2021).

## 2020

1. Xuan-Phi Nguyen, **Shafiq Joty**, Wu Kui, and Aw Ai Ti., “[Data Diversification: A Simple Strategy for Neural Machine Translation.](#)”, *In 2020 Conference on Neural Information Processing Systems (NeurIPS’20)*, Vancouver, Canada (2020).
2. Jiuxiang Gu, Jason Kuen, **Shafiq Joty**, Jianfei Cai, Vlad Morariu, Handong Zhao, and Tong Sun., “[Self-Supervised Relationship Probing.](#)”, *In 2020 Conference on Neural Information Processing Systems (NeurIPS’20)*, Vancouver, Canada (2020).
3. Xuan-Phi Nguyen, **Shafiq Joty**, Steven Hoi, and Richard Socher., “[Tree-Structured Attention with Hierarchical Accumulation.](#)”, *In International Conference on Learning Representations (ICLR-2020)*, Addis Ababa, Ethiopia (2020).
4. Samson Tan, **Shafiq Joty**, Min-Yen Kan, and Richard Socher., “[It’s Morphin’ Time! Combating Linguistic Discrimination with Inflectional Perturbations.](#)”, *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020)*, Seattle, USA (2020).



5. Yifan Gao, Chien-Sheng Wu, **Shafiq Joty**, Caiming Xiong, Richard Socher, Irwin King, Michael Lyu, and Steven C.H., “[Explicit Memory Tracker with Coarse-to-Fine Reasoning for Conversational Machine Reading](#).”, In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020)*, Seattle, USA (2020).
6. Thanh-Tung Nguyen, Xuan-Phi Nguyen, **Shafiq Joty**, and Xiaoli Li., “[Efficient Constituency Parsing by Pointing](#).”, In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020)*, Seattle, USA (2020).
7. Thanh-Tung Nguyen, Xuan-Phi Nguyen, **Shafiq Joty**, and Xiaoli Li., “[Differentiable Window for Dynamic Local Attention](#).”, In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020)*, Seattle, USA (2020).
8. Tasnim Mohiuddin, M Saiful Bari, and **Shafiq Joty**, “[LNMap: Departures from Isomorphic Assumption in Bilingual Lexicon Induction Through Non-Linear Mapping in Latent Space](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
9. Prathyusha Jwalapuram, **Shafiq Joty** and Youlin Shen, “[Pronoun-Targeted Finetuning for NMT with Hybrid Losses](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
10. Samson Tan, **Shafiq Joty**, Lav R., and Min-Yen Kan., “[Mind Your Inflections! Improving NLP for Non-Standard English with Base-Inflection Encoding](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
11. Weishi Wang, **Shafiq Joty**, Steven C.H., “[Response Selection for Multi-Party Conversations with Dynamic Topic Tracking](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
12. Tao Yu and **Shafiq Joty**, “[Online Conversation Disentanglement with Pointer Networks](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
13. Yue Wang, **Shafiq Joty**, Michael R., Irwin King, Caiming Xiong, and Steven C.H., “[VD-BERT: A Unified Vision and Dialog Transformer with BERT](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
14. Yifan Gao, Chien-Sheng Wu, Jingjing Li, **Shafiq Joty**, Steven C.H., Caiming Xiong, Irwin King, and Michael Lyu. , “[Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).
15. Bosheng Ding, Linlin Liu, Lidong Bing, Canasai Kruengkrai, Thien Hai, **Shafiq Joty**, Luo Si, and Chunyan Miao., “[DAGA: Data Augmentation with a Generation Approach for Low-resource Tagging Tasks](#).”, In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2020)*, Punta Cana, Dominican Republic (2020).

16. Xiangxi Shi, Xu Yang, Jiuxiang Gu, **Shafiq Joty**, and Jianfei Cai, “[Finding It at Another Side: A Viewpoint-Adapted Matching Encoder for Change Captioning](#)”, In *Proceedings of the European Conference on Computer Vision (ECCV’20)*, Online (2020).
17. Saiful Bari, **Shafiq Joty**, and Prathyusha Jwalapuram., “[Zero-Resource Cross-Lingual Named Entity Recognition](#).”, In *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020)*, New York, USA (2020).
18. Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, **Shafiq Joty**, Eng Siong, and Bin Ma., “[Universal Speech Transformer](#).”, In *21st Annual Conference of the International Speech Communication Association (Interspeech-2020)*, Shanghai, China (2020).
19. Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, **Shafiq Joty**, Eng Siong, and Bin Ma., “[Speech Transformer with Speaker Aware Persistent Memory](#).”, In *21st Annual Conference of the International Speech Communication Association (Interspeech-2020)*, Shanghai, China (2020).
20. Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, **Shafiq Joty**, Eng Siong, and Bin Ma., “[Cross Attention with Monotonic Alignment for Speech Transformer](#).”, In *21st Annual Conference of the International Speech Communication Association (Interspeech-2020)*, Shanghai, China (2020).

## 2019

1. Linlin Liu\*, Xiang Lin\*, **Shafiq Joty**, Simeng Han, and Lidong Bing., “[Hierarchical Pointer Net Parsing](#).”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2019)*, pages x-x, Hong Kong, China (2019).
2. Han-Cheol Moon\*, Tasnim Mohiuddin\*, **Shafiq Joty\***, and Chi Xu., “[A Unified Neural Coherence Model](#).”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2019)*, pages x-x, Hong Kong, China (2019).
3. Prathyusha Jwalapuram, **Shafiq Joty**, Irina Temnikova, and Preslav Nakov., “[Evaluating Pronominal Anaphora in Machine Translation: An Evaluation Measure and a Test Suite](#).”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2019)*, pages x-x, Hong Kong, China (2019).
4. Swaraj Khadanga, Karan Aggarwal, **Shafiq Joty**, and Jaideep Srivastava., “[Using Clinical Notes with Multimodal Learning for ICU Management](#).”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2019)*, pages x-x, Hong Kong, China (2019).
5. Jiuxiang Gu, **Shafiq Joty**, Jianfei Cai, Handong Zhao, Xu Yang, and Gang Wang, “[Unpaired Image Captioning via Scene Graph Alignments](#)”, In *Proceedings of the International Conference on Computer Vision (ICCV’19)*, Seoul, Korea (2019).
6. Xiangxi Shi, Jianfei Cai, **Shafiq Joty**, and Jiuxiang Gu, “[Watch It Twice: Video Captioning with a Refocused Video Encoder](#)”, In *Proceedings of the 27th ACM International Conference on Multimedia (ACMMM’19)*, Nice, France (2019).
7. Xiang Lin\*, **Shafiq Joty\***, Prathyusha Jwalapuram and M Saiful Bari, “[A Unified Linear-Time Framework for Sentence-Level Discourse Parsing](#)”, In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL’19)*, Florence, Italy (2019).

8. Jing Ma, Wei Gao, **Shafiq Joty**, and Kam-Fai Wong, “[Sentence-Level Evidence Embedding for End-to-End Claim Verification with Hierarchical Attention Networks](#)”, In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL’19)*, Florence, Italy (2019).
9. **Shafiq Joty**, Giuseppe Carenini, Raymond Ng, and Gabriel Murray., “[Discourse Processing and Its Applications](#)”, In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Tutorial (ACL-2019 Tutorial)*, Florence, Italy (2019).
10. Tasnim Mohiuddin and **Shafiq Joty**, “[Revisiting Adversarial Autoencoder for Unsupervised Word Translation with Cycle Consistency and Improved Training](#)”, In *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-19)*, Minneapolis, USA (2019).
11. Tasnim Mohiuddin, Thanh-Tung Nguyen and **Shafiq Joty** (Equal Contributions), “[Adaptation of Hierarchical Structured Models for Speech Act Recognition in Asynchronous Conversation](#)”, In *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-19)*, Minneapolis, USA (2019).
12. Karan Aggarwal, **Shafiq Joty**, Luis Fernandez-Luque, and Jaideep Srivastava., “[Adversarial Unsupervised Representation Learning for Activity Time-Series](#)”, In *Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-2019)*, Honolulu, Hawaii (2019).
13. Sameer Khurana, **Shafiq Joty**, Ahmed Ali, and James Glass, “[A Factorial Deep Markov Model For Unsupervised Disentangled Representation Learning From Speech](#)”, In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2019)*, Brighton, UK (2019).
14. Siyuan Liu, Sourav S Bhowmick, Wanlu Zhang, Shu Wang, Wanyi Huang, and **Shafiq Joty**, “[NEURON: Query Execution Plan Meets Natural Language Processing For Augmenting DB Education](#)”, In *Proceedings of 45th ACM SIGMOD International Conference on Management of Data (SIGMOD-2019 Demo)*, Amsterdam, NL (2019).

## 2018

1. **Shafiq Joty**, Giuseppe Carenini, Raymond Ng, and Gabriel Murray., “[Discourse Processing and Its Applications in Text Mining](#)”, In *18th IEEE International Conference on Data Mining: Tutorial (ICDM-2018 Tutorial)*, Singapore (2018).
2. Gabriel Murray, **Shafiq Joty**, and Giuseppe Carenini., “[NLP for Conversations: Sentiment, Summarization, and Group Dynamics.](#)”, In *Proceedings of the 27th International Conference on Computational Linguistics: Tutorial Abstracts (COLING-2018 Tutorial)*, pages 1-4, Santa Fe, New Mexico, USA (2018).
3. **Shafiq Joty**, Lluís Màrquez and Preslav Nakov., “[Joint Multitask Learning for Community Question Answering Using Task-Specific Embeddings.](#)”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2018)*, pages x-x, Brussels, Belgium (2018).
4. **Shafiq Joty**, Tasnim Mohiuddin, and Dat Nguyen (equal contributions), “[Coherence Modeling of Asynchronous Conversations: A Neural Entity Grid Approach](#)”, In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL’18)*, Melbourne, Australia (2018).

5. Firoj Alam, **Shafiq Joty**, and Muhammad Imran., “[Domain Adaptation with Adversarial Training and Graph Embeddings](#)”, *In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL’18)*, Melbourne, Australia (2018).
6. Jiuxiang Gu, **Shafiq Joty**, Jianfei Cai, and Gang Wang, “[Unpaired Image Captioning by Language Pivoting](#)”, *In Proceedings of the European Conference on Computer Vision (ECCV-2018)*, Munich, Germany (2018).
7. Qing Li, Qingyi Tao, **Shafiq Joty**, Jianfei Cai and Jiebo Luo, “[VQA-E: Explaining, Elaborating, and Enhancing Your Answers for Visual Questions](#)”, *In Proceedings of the VQA-E: Explaining, Elaborating, and Enhancing Your Answers for Visual Questions (ECCV-2018)*, Munich, Germany (2018).
8. Jiuxiang Gu, Jianfei Cai, **Shafiq Joty**, Li Niu, and Gang Wang, “[Look, Imagine and Match: Improving Textual-Visual Cross-Modal Retrieval with Generative Models](#)”, *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2018)*, Salt Lake City, Utah, United States (2018).
9. Chin Yao, Kaiqi Zhao, **Shafiq Joty**, and Gao Cong, “[Aspect-based Neural Recommender](#)”, *In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM-2018)*, Turin, Italy (2018).
10. Jing Li, Aixin Sun, and **Shafiq Joty**, “[SegBot: A Generic Neural Text Segmentation Model with Pointer Network](#)”, *In Proceedings of the 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence (IJCAI-ECAI-2018)*, Stockholm, Sweden (2018).
11. Karan Aggarwal, Swaraj Khadanga, **Shafiq Joty**, Louis Kazaglis, and Jaideep Srivastava, “[A Structured Learning Approach with Neural Conditional Random Fields for Sleep Staging](#)”, *In Proceedings of the IEEE Big Data Conference 2018 (IEEE Big Data-2018)*, Seattle, USA (2018).
12. Muhammad Ebraheem, Saravanan Thirumuruganathan, **Shafiq Joty**, Mourad Ouzzani and Nan Tang, “[Distributed Representations of Tuples for Entity Resolution](#)”, *In Proceedings of the Forty-fourth International Conference on Very Large Data Bases (VLDB-2018)*, Rio de Janeiro, Brazil (2018).
13. Firoj Alam, **Shafiq Joty**, and Muhammad Imran, “[Graph Based Semi-supervised Learning with Convolutional Neural Networks to Classify Crisis Related Tweets](#)”, *In Proceedings of the Twelfth International Conference on Web and Social Media (ICWSM-2018)*, Stanford, United States (2018).

## 2017

1. Tanay Kumar Saha, **Shafiq Joty** and Mohammad Al Hasan., “[CON-S2V: A Generic Framework for Incorporating Extra-Sentential Context into Sen2Vec.](#)”, *In Proceedings of the European Conference on Machine Learning and practice of knowledge discovery in databases (ECML-PKDD-2017)*, Macedonia, Skopje (2017).
2. Tanay Kumar Saha, **Shafiq Joty**, Naeemul Hassan, and Mohammad Al Hasan., “[Regularized and Retrofitted models for Learning Sentence Representation with Context.](#)”, *In Proceedings of the 26th ACM International Conference on Information and Knowledge Management (CIKM-2017)*, Singapore (2017).

3. Dat Tien Nguyen and **Shafiq Joty** (co-lead author), “[A Neural Local Coherence Model](#).”, *In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL-2017)*, pages 1320-1330, Vancouver, Canada (2017).
4. **Shafiq Joty**, Preslav Nakov and Lluís Màrquez, “[Cross-language Learning with Adversarial Neural Networks](#).”, *In Proceedings of The SIGNLL Conference on Computational Natural Language Learning (CoNLL-2017)*, pages 226-237, Vancouver, Canada (2017).
5. Enamul Hoque, **Shafiq Joty**, Lluís Màrquez, and Giuseppe Carenini., “[CQAVis: Visual Text Analytics for Community Question Answering](#).”, *In Proceedings of the 22nd annual meeting of the intelligent user interfaces (IUI-2017)*, pages 161-172, Limassol, Cyprus (2017).
6. Giovanni Da San Martino, Salvatore Romeo, Alberto Barron-Cedeno, Shafiq Joty, Lluís Màrquez, Alessandro Moschitti and Preslav Nakov., “[Cross-Language Question Re-ranking](#).”, *In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-2017)*, pages 1145-1148, Tokyo, Japan (2017).
7. Dat Tien Nguyen, Kamla Al-Mannai, **Shafiq Joty**, Hassan Sajjad, Muhammad Imran and Prasenjit Mitra., “[Robust Classification of Crisis-Related Data on Social Networks using Convolutional Neural Networks](#).”, *In Proceedings of the 11-th International AAAI Conference on Weblogs and Social Media (ICWSM-2017)*, pages 632-635, Montreal, Canada (2017).

## 2016

1. **Shafiq Joty** and Enamul Hoque., “[Speech Act Modeling of Written Asynchronous Conversations with Task-Specific Embeddings and Conditional Structured Models](#).”, *In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL-2016)*, pages 1746-1756, Berlin, Germany (2016).
2. **Shafiq Joty**, Lluís Màrquez and Preslav Nakov., “[Joint Learning with Global Inference for Comment Classification in Community Question Answering](#).”, *In Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-2016)*, pages 703-713, San Diego, USA (2016).
3. Nadir Durrani, Hassan Sajjad, **Shafiq Joty**, and Ahmed Abdelali., “[A Deep Fusion Model for Domain Adaptation in Phrase-based MT](#).”, *In Proceedings of the 26th International Conference on Computational Linguistics (COLING-2016)*, pages 3177–3187, Osaka, Japan (2016).
4. Enamul Hoque, **Shafiq Joty**, Lluís Màrquez, Alberto Barrón-Cedeño, Alessandro Moschitti, Giovanni Da San Martino, Preslav Nakov, Salvatore Romeo and Giuseppe Carenini., “[An Interactive System for Exploring Community Question Answering Forums](#).”, *In Proceedings of the 26th International Conference on Computational Linguistics (COLING-2016)*, pages 1-5, Osaka, Japan (2016).

## 2015

1. **Shafiq Joty**, Hassan Sajjad, Nadir Durrani, Kamla Al-Mannai, Ahmed Abdelali and Stephan Vogel., “[How to Avoid Unwanted Pregnancies: Domain Adaptation using Neural Network Models](#).”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2015)*, pages 1259-1270, Lisbon, Portugal (2015).



2. **Shafiq Joty**, Alberto Barrón-Cedeño, Giovanni Da San Martino, Alessandro Moschitti, Lluís Màrquez and Preslav Nakov., “[Global Thread-level Inference for Comment Classification in Community Question Answering.](#)”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2015)*, pages 573-578, Lisbon, Portugal (2015).
3. Pengfei Liu, **Shafiq Joty**, and Helen Meng., “[Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings.](#)”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2015)*, pages 1433-1443, Lisbon, Portugal (2015).
4. Francisco Guzmán, **Shafiq Joty**, Lluís Màrquez, and Preslav Nakov., “[Pairwise Neural Machine Translation Evaluation.](#)”, *In Proceedings of the 53rd annual meeting of the Association for Computational Linguistics (ACL-2015)*, pages 805 -814, Beijing, China (2015).
5. Alberto Barrón-Cedeño, Simone Filice, Giovanni Da San Martino, **Shafiq Joty**, Lluís Màrquez, Preslav Nakov and Alessandro Moschitti., “[Thread-level Information for Comment Classification in Community Question Answering.](#)”, *In Proceedings of the 53rd annual meeting of the Association for Computational Linguistics (ACL-2015)*, pages 687-693, Beijing, China (2015).
6. Nadir Durrani, Hassan Sajjad, **Shafiq Joty**, Ahmed Abdelali and Stephan Vogel., “[Using Joint Models for Domain Adaptation in Statistical Machine Translation.](#)”, *In Proceedings of the Association for Machine Translation in the Americas (AMTA-2015)*, Miami, USA (2015).

## 2014

1. **Shafiq Joty**, Alessandro Moschitti., “[Discriminative Reranking of Discourse Parses Using Tree Kernels.](#)”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, pages 2049-2060, Doha, Qatar (2014).
2. Francisco Guzmán, **Shafiq Joty**, Lluís Màrquez, Alessandro Moschitti, Preslav Nakov and Massimo Nicosia., “[Learning to Differentiate Better from Worse Translations.](#)”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, pages 214-220, Doha, Qatar (2014).
3. Iman Saleh, Alessandro Moschitti, Preslav Nakov, Lluís Màrquez and **Shafiq Joty.**, “[Semantic Kernels for Semantic Parsing.](#)”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, pages 436-442, Doha, Qatar (2014).
4. Francisco Guzmán, **Shafiq Joty**, Lluís Màrquez, and Preslav Nakov., “[Using Discourse Structure Improves Machine Translation Evaluation.](#)”, *In Proceedings of the 52nd annual meeting of the Association for Computational Linguistics (ACL-2014)*, pages 687-698, Baltimore, USA (2014).
5. Iman Saleh, **Shafiq Joty**, Lluís Màrquez, Alessandro Moschitti, Preslav Nakov, Scott Cyphers and Jim Glass., “[A Study of Using Syntactic and Semantic Structures for Concept Segmentation and Labeling.](#)”, *In Proceedings of the 25th International Conference on Computational Linguistics (COLING-2014)*, pages 193-202, Dublin, Ireland (2014).

## 2013

1. **Shafiq Joty**, Giuseppe Carenini, Raymond Ng and Yashar Mehdad., “[Combining Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis.](#)”, In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. (ACL-2013)*, pages 486-496, Sofia, Bulgaria (2013).
2. Yashar Mehdad, Giuseppe Carenini, Raymond Ng and **Shafiq Joty.**, “[Towards Topic Labeling with Phrase Entailment and Aggregation.](#)”, In *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. (NAACL-2013)*, pages 179-189, Atlanta, USA (2013).
3. Maryam Tavafi, Yashar Mehdad, **Shafiq Joty**, Giuseppe Carenini, Raymond Ng., “[Dialogue Act Recognition in Synchronous and Asynchronous Conversations.](#)”, In *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue. (SIGDIAL-2013)*, pages 117-121, Metz, France. (2013).

## 2012

1. **Shafiq Joty**, Giuseppe Carenini and Raymond Ng., “[A Novel Discriminative Framework for Sentence-Level Discourse Analysis.](#)”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing and the Conference on Natural Language Learning. (EMNLP-2012)*, pages 904-915, Jeju, Korea (2012).

## 2011

1. **Shafiq Joty**, Giuseppe Carenini and Chin-Yew Lin., “[Unsupervised Modeling of Dialog Acts in Asynchronous Conversations.](#)”, In *Proceedings of the twenty second International Joint Conference on Artificial Intelligence. (IJCAI-2011)*, pages 1807-1813, Barcelona, Spain (2011).
2. **Shafiq Joty**, Giuseppe Carenini, Gabriel Murray and Raymond Ng., “[Supervised Topic Segmentation of Email Conversations.](#)”, In *Fifth International AAAI Conference on Weblogs and Social Media. (ICWSM-2011)*, pages 530-533, Barcelona, Spain (2011).
3. Nicholas FitzGerald, Giuseppe Carenini, Gabriel Murray and **Shafiq Joty.**, “[Exploiting Conversational Features to Detect High-Quality Blog Comments.](#)”, In *Proceedings of the Canadian Conference on Artificial Intelligence (CAI-2011)*, pages 122-127, St. Johns, Newfoundland (2011).

## 2010

1. **Shafiq Joty**, Giuseppe Carenini, Gabriel Murray and Raymond Ng., “[Exploiting Conversation Structure in Unsupervised Topic Segmentation for Emails.](#)”, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing and the Conference on Natural Language Learning. (EMNLP-2010)*, pages 388-398, Massachusetts, USA. (2010).

## 2009

1. Yllias Chali, Sadid Hasan and **Shafiq Joty.**, “[Do Automatic Annotation Techniques Have Any Impact on Supervised Complex Question Answering?](#)”, In *Proceedings of the ACL-IJCNLP 2009 Conference. (ACL-2009)*, pages 329-332, Suntec, Singapore. (2009).
2. Yllias Chali, Sadid Hasan and **Shafiq Joty.**, “[A SVM-Based Ensemble Approach to Multi-Document Summarization.](#)”, In *Proceedings of the 22nd Canadian Conference on Artificial Intelligence (CAI-2009)*, pages 329-332, Kelowna, Canada. (2009).



3. Yllias Chali, Sadid Hasan and **Shafiq Joty**., “[Supervised Approaches to Complex Question Answering](#).”, *In Proceedings of Conference of the Pacific Association for Computational Linguistics (PACLING-2009)*, pages 121-126, Sapporo, Japan. (2009).

## 2008

1. Yllias Chali and **Shafiq Joty**., “[Selecting Sentences for Answering Complex Questions](#).”, *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2008)*, pages 304-313, Hawaii, USA. (2008). (I was a co-lead author.)
2. Yllias Chali and **Shafiq Joty**., “[Improving the Performance of the Random Walk Model for Answering Complex Questions](#).”, *In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-2008)*, pages 9-12, OH, USA. (2008). (I was a co-lead author.)
3. Yllias Chali and **Shafiq Joty**., “[Exploiting Syntactic and Shallow Semantic Kernels to Improve Random Walks for Complex Question Answering](#).”, *In Proceedings of the 20th IEEE Int’l Conference on Tools with Artificial Intelligence (ICTAI-2008)*, pages 123-130, OH, USA. (2008). (I was a co-lead author.)
4. Yllias Chali and **Shafiq Joty**., “[Answering Complex Questions Using Query-focused Summarization Technique](#).”, *In Proceedings of the 20th IEEE Int’l Conference on Tools with Artificial Intelligence (ICTAI-2008)*, pages 131-134, OH, USA. (2008). (I was a co-lead author.)
5. Yllias Chali and **Shafiq Joty**., “[Unsupervised Approach for Selecting Sentences in Query-based Summarization](#).”, *In Proceedings of the 21st FLAIRS conference (FLAIRS-2007)*, pages 47-52, Florida, USA. (2008). (I was a co-lead author.)
6. Gabriel Murray, **Shafiq Joty**, Giuseppe Carenini and Raymond Ng., “[The University of British Columbia at TAC 2008](#).”, *In Proceedings of the Text Analysis Conference (TAC-2008)*, Rochester, USA. (2008).
7. Yllias Chali and **Shafiq Joty** and Sadid-Al-Hasan., “[UofL at TAC 2008 Update Summarization and Question Answering](#).”, *In Proceedings of the Text Analysis Conference (TAC-2008)*, Rochester, USA. (2008).

## 2007

1. Yllias Chali and **Shafiq Joty**., “[University of Lethbridge’s Participation in DUC-2007 Main Task](#).”, *In Proceedings of the Document Understanding Conference (DUC-2007)*, Rochester, USA. (2007). (I was a co-lead author.)
2. Yllias Chali and **Shafiq Joty**., “[University of Lethbridge’s Participation in TREC-2007 Main Task](#).”, *In Proceedings of the Sixteenth Text REtrieval Conference (TREC-2007)*, Gaithersburg, USA. (2007). (I was a co-lead author.)
3. **Shafiq Joty** and Sadid Hasan., “[Advances in Focused Retrieval: A General Review](#).”, *In Proceedings of the 10th IEEE International Conference on Computer and Information Technology (ICCIT-2007)*, pages 1-5, Dhaka, Bangladesh. (2007).

## PEER-REVIEWED WORKSHOP PAPERS

---

1. Nafise Sadat Moosavi, Angela Fan, Vered Shwartz, Goran Glavas, **Shafiq Joty**, Alex Wang, and Thomas Wolf., “[Proceedings of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing](#)”, In *Proceedings of SustaiNLP Workshop (EMNLP’20 SustaiNLP Workshop)*, Online (2020).
2. Dat Nguyen, **Shafiq Joty**, Basma Boussaha, and Maarten de Rijke., “[Thread Reconstruction in Conversational Data using Neural Coherence Models.](#)”, In *Proceedings of the Neu-IR 2017 SIGIR Workshop on Neural Information Retrieval (NeuIR’17)*, Tokyo, Japan (2017).
3. Alberto Barrón-Cedeño, Giovanni Da San Martino, **Shafiq Joty**, Alessandro Moschitti, Fahad Al-Obaidli, Salvatore Romeo, Kateryna Tymoshenko and Antonio Uva., “[ConvKN at SemEval-2016 Task 3: Answer and Question Selection for Question Answering on Arabic and English Fora.](#)”, In *Proceedings of the 10th International Workshop on Semantic Evaluation. (SemEval 2016)*, San Diego, USA (2016). (**2<sup>nd</sup> place in both English & Arabic subtasks of Task 3**)
4. Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Alessandro Moschitti, **Shafiq Joty**, Fahad Al-Obaidli, Kateryna Tymoshenko, Antonio Uva and Daniele Bonadiman., “[Addressing Community Question Answering in English and Arabic.](#)”, In *Second WebQA workshop. (SIGIR-WebQA 2016)*, Pisa, Italy (2016).
5. Dat Tien Nguyen, **Shafiq Joty**, Muhammad Imran, Hassan Sajjad, and Prasenjit Mitra., “[Applications of Online Deep Learning for Crisis Response Using Social Media Information.](#)”, In *the 4th International Workshop on Social Web for Disaster Management (co-located with CIKM 2016), (SWDM 2016)*, Indianapolis, USA (2016).
6. Massimo Nicosia, Simone Filice, Alberto Barrón-Cedeño, Iman Saleh, Hamdy Mubarak, Wei Gao, Preslav Nakov, Giovanni Da San Martino, Alessandro Moschitti, Kareem Darwish, Lluís Màrquez, **Shafiq Joty** and Walid Magdy., “[QCRI: Answer Selection for Community Question Answering - Experiments for Arabic and English.](#)”, In *Proceedings of the 9th International Workshop on Semantic Evaluation. (SemEval-2015)*, pages 203-209, Denver, USA (2015). (**Winner in Arabic and 3<sup>rd</sup> place in English subtask of Task 3**)
7. **Shafiq Joty**, Francisco Guzman, Lluís Marquez and Preslav Nakov., “[DiscoTK: Using Discourse Structure for Machine Translation Evaluation.](#)”, In *Proceedings of the Ninth Workshop on Statistical Machine Translation (WMT-2014)*, pages 402-408, Baltimore, USA (2014). (**Winner of the metric task.**)
8. Enamul Hoque, Giuseppe Carenini and **Shafiq Joty.**, “[Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System.](#)”, In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces (ILLVI-2014)*, pages 45-52, Baltimore, USA. (2014).
9. Wei Jin, **Shafiq Joty**, Giuseppe Carenini and Raymond Ng., “[Detecting Informative Blog Comments using Tree Structured Conditional Random Fields.](#)”, *The Pacific Northwest Regional NLP Workshop (NWNLP-2012)*, Microsoft Research, Redmond, USA. (2012).
10. **Shafiq Joty**, Giuseppe Carenini, Gabriel Murray and Raymond Ng., “[Finding Topics in Emails: Is LDA enough?](#)”, *NIPS-2009 workshop on applications for topic models: text and beyond. (NIPS-2009)*, Whistler, Canada (2009).

11. Yllias Chali and **Shafiq Joty**., “[Word Sense Disambiguation Using Lexical Cohesion](#).”, *In Proceedings of the 4th International Conference on Semantic Evaluations (SemEval-2007)*, pages 476-479, Prague, Czech Republic. (2007).

## RESEARCH GRANTS

---

1. NLP for WHO’s Workforce Intelligence., “WHO International”, 375,000 USD (Dec, 2021 – Oct, 2022).
2. Trustworthy and Explainable AI: Text Data Augmentation, Multi-Lingual NLP Generation., “AISG2-PhD-2021-01-001”, 480,000 SGD (Jan, 2021 – Sep, 2025).
3. Inference Capable Natural Language Generation for Conversational AI., “AI.SG 100ER”, 400,000 SGD (Sep, 2021 – Aug, 2023).
4. Zero-shot and few-shot learning for NER with/without a small In-domain KB., “Alibaba-NTU Singapore Joint Research Institute”, 136,104 SGD (Nov, 2020 – Nov, 2021).
5. AI-supported collaborative activity design., “AI.SG 100 Experiments”, 180,665.76 SGD (Apr, 2020 – Sep, 2021).
6. Customised Tamil-English Translation Engine., “Joint Project with A-star”, 149,400 SGD (Sep, 2019 – Apr, 2021).
7. Knowledge Graph Mining for Educational Domain., “Delta-NTU-NRF”, 199,964 SGD (Jan, 2020 – Dec, 2021).
8. Language Understanding and Generation with Deep Neural Networks., “Start Up-Grant (NTU, Singapore)”, 200,000 SGD (Jul, 2017 – Jul, 2020).
9. Multilingual Language Processing Using Generative Adversarial Autoencoders., “Tier-1 program (NTU, Singapore)”, 100,000 SGD (Mar, 2018 - Aug, 2019).
10. A Malay-English Translation System., “SCSE demo fund (NTU, Singapore)”, (Jan, 2018 – Jun, 2018).

## TUTORIALS AT INTERNATIONAL CONFERENCES

---

1. [Discourse Analysis and Its Applications](#), **Shafiq Joty**, Giuseppe Carenini, Raymond Ng, and Gabriel Murray, **ACL-2019**, Florence, Italy. ([Slides](#))
2. [Discourse Processing and Its Applications in Text Mining](#), **Shafiq Joty**, Giuseppe Carenini, Raymond Ng, and Gabriel Murray, **ICDM-2018**, Singapore. ([Slides](#))
3. [NLP for Conversations: Sentiment, Summarization, and Group Dynamics](#), Gabriel Murray, **Shafiq Joty**, and Giuseppe Carenini., **COLING-2018**, Santa Fe, New Mexico. ([Slides](#))

## INVITED TALKS

---

1. “Model, Task and Data Engineering for NLP”, UBC, Canada, June. 2021
2. “Model, Task and Data Engineering for NLP”, York University, Canada, June. 2021
3. [“Model, Task and Data Engineering for NLP”](#), NUS, Singapore, June. 2021
4. “Towards Robust Multi-lingual Multi-dialectal NLP Models”, Alibaba Damo Academy, Singapore, Mar. 2021
5. “NLP Research at NTU-NLP”, Alibaba NTU JRI, Online, Nov. 2020
6. “Discourse Analysis with Deep Learning”, TU-Darmstadt, Germany, July. 2019
7. “Structured Models for Discourse Analysis and NLP Applications ”, Alibaba, Singapore, Jan. 2019
8. [“Deep Structured Models for Natural Language Processing”](#), NTU-SenseTime Research Seminar, NTU, Singapore, Nov. 2018
9. “Deep Structured Models for Natural Language Processing ”, Salesforce AI Research, Palo Alto, USA, Oct. 2018
10. “Deep Learning for Multilingual and Multimodal Processing”, DSO National Laboratories, Singapore, Oct. 2018
11. [“Deep Learning for Multilingual and Multimodal Processing”](#), Industry-NTU Dialogue: AI Research, Innovation and Talent, NTU, Singapore, Sep. 2018
12. “Discourse Analysis and Its Applications”, Gwangju Institute of Science and Technology (GIST), South Korea, Dec. 2017
13. “How Computational Linguistics can help Computational History”, Johannesburg Institute for Advanced Study (JIAS), Johannesburg, South Africa, Nov. 2017
14. “Discourse Analysis and Its Applications”, National University of Singapore (NUS), Singapore, Oct. 2017
15. “Discourse Analysis and Its Applications”, Institute for Infocomm Research A\*Star, Singapore, Sep. 2017
16. [“Structured Predictions for Discourse Analysis and Applications in Natural Language Processing”](#), Singapore University of Technology and Design (SUTD), Singapore, Sep. 2017
17. “Structured Predictions for Discourse Analysis and Applications in Natural Language Processing”, University of Calgary, Calgary, Canada, Mar. 2017
18. “Structured Predictions for Discourse Analysis and Applications in Natural Language Processing”, University of Waterloo, Waterloo, Canada, Feb. 2017
19. “Structured Predictions for Discourse Analysis and Applications in Natural Language Processing”, Nanyang Technological University (NTU), Singapore, Nov. 2016
20. “Introduction to Machine Learning”, Global Entrepreneurship Week, Doha, Nov. 2015

21. “Discourse Analysis & its Applications”, Chinese University of Hong Kong., Oct. 2015
22. “CODRA: A Complete Discriminative Framework for Rhetorical Analysis”, Universite Paul Sabatier, Toulouse, France., Sep. 2014
23. “Discourse Analysis of Asynchronous Conversations”, University of British Columbia, Vancouver, Canada., Dec. 2013
24. “Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis”, NSERC Business Intelligence Network AGM, IBM Software Labs, Canada., Nov. 2012
25. “A Novel Discriminative Framework for Sentence-Level Discourse Analysis”, NW-NLP workshop, Microsoft Research Redmond., May, 2012
26. “Discourse Analysis of Asynchronous Conversations (e.g., emails, blogs)”, Business Intelligence (BI) Workshop-NSERC BIN Watson, IBM Watson center, Feb, 2012
27. “Discourse Analysis of Asynchronous Conversations (e.g., emails, blogs)”, UBC Social Network Group, Feb, 2012
28. “Unsupervised Modeling of Dialog Acts in Asynchronous Conversations”, UBC Laboratory for Computational Intelligence (LCI) Forum Talk, Apr, 2011
29. “Topic and Dialog Act Modeling of Asynchronous Conversations”, Microsoft Research Asia, Nov, 2010
30. “What are they talking about?” Finding Topics in Email Conversations”, UBC Laboratory for Computational Intelligence Forum (LCI) Talk, Feb, 2010
31. “Answering Simple and Complex Questions”, University of Alberta NLP Group, Jan, 2008

## SERVICE

---

1. Senior area chair for [ACL-2022](#) (ML track), [EMNLP-2021](#) (Applications track), Area chair for [ACL-2021](#) (ML track), [ACL-2020](#) (Discourse track), [NAACL-2021](#) (Discourse track), [ACL-2019](#) (QA), [EMNLP-2019](#) (Discourse track)
2. Action editor for [ACL rolling review](#)
3. Associate editor for ACM Transactions on Asian and Low Resource Language Resource language Processing (TALLIP) [Jan 2020, Feb 2021]
4. Senior Program Committee member for [IJCAI-2019](#), AAAI-2020
5. Grant Proposal Reviewing: [NSERC Canada Discovery Grant](#) [2018].
6. Reviewing: TACL [2015-2019], ACL/EMNLP/SigDial/\*SEM/NAACL/IJCAI/IUI/KDD.
7. Associate VP position in the Bangladesh Students’ Association, UBC (2009-2010).
8. Organizing committee member of EMNLP-2014, ICCIT-2005.

## PROFESSIONAL MEMBERSHIP

---

- ACL, SIGDAT, SIGGEN, SIGSEM, ACM, IEEE.

## STUDENTS

---

### • Current PhD Students

1. [Esther \(Ruochen\) Zhao](#) [Jan, 2021 - Present, AISG Fellowship]
2. [Mathieu Ravaut](#) [Jan, 2021 - Present, A\*Star RSS-SINGA fellowship]
3. [Chen Hailin](#) [Jan, 2021 - Present, Salesforce IPP NTU]
4. [Chengwei Qin](#) [Jan, 2021 - Present, NTU Graduate Scholarship]
5. [Xuan Phi Nguyen](#) [Aug, 2019 - Present, A\*STAR ACIS fellowship] [SDSC dissertation fellowship](#)
6. [Samson Tan \(industry advisor\)](#) [Aug, 2019 - Present, Salesforce IPP NUS] [NUS Dean's Research Excellence Award 2021, Research Achievement Award, 2020](#)
7. [Lin Xiang](#) [Jan, 2019 - Present, Self-funded (RA)]
8. [Han-Cheol Moon](#) [Jan, 2019 - Present, SINGA-SINGA fellowship]
9. [Saiful Bari](#) [Jan, 2019 - Present, NTU Graduate Scholarship]
10. [Prathyusha Jwalapuram](#) [Aug, 2018 - Present, NTU Graduate Scholarship]
11. [Tasnim Mohiuddin](#) [Jan, 2018 - Present, NTU Graduate Scholarship]
12. [Linlin Liu \(co-supervision\)](#) [Jan, 2019 - Present]
13. [Bosheng Ding \(co-supervision\)](#) [Jan, 2020 - Present]

### • Graduated Students (PhD)

1. [Nguyen Thanh Tung](#) [Current Position: Research Scientist, I2R A-star, Singapore]  
**PhD Thesis:** Pointing Network for Natural Language Understanding

### • Undergraduate FYP Students

1. [Han Simeng](#) [Mar, 2019 - Jul, 2021] [[Best Final Year Thesis Gold Medal](#); Current Position: PhD student in CS @ Yale University, USA]
2. [Shen Youlin](#) [Mar, 2020 - Dec, 2020]
3. [Rixie Tiffany Ko Leong](#) [Jan, 2021 - Dec, 2021]
4. [Angshuk Dutta](#) [Aug, 2021 - ]

## • Visiting Students

1. [Sameer Khurana](#) [May, 2018 - Aug, 2018]. (Now PhD student at MIT-CSAIL, USA)
2. [Jing MA](#) [Dec, 2018 - Aug, 2019]. (Now Asst Prof at Hong Kong Baptist University)
3. Lingzhi Wang [Aug 2021 - Nov 2021] (Now a PhD student at HKUST, Hong Kong)
4. [Enamul Hoque](#) [Feb, 2016 - May, 2016] (Now Asst. Prof. at York Univ, Canada)
5. [Aarti Sathyanarayana](#) [July, 2016 - Dec, 2016] (Now Asst Prof at Northeastern University, USA)

## REFERENCES

---

### [Dr. Giuseppe Carenini](#)

Professor  
Co-Director of Data Science Institute  
Dept. of Computer Science  
University of British Columbia  
2366 Main Mall, Vancouver, B.C.  
Canada V6T 1Z4  
☎ +1 (604) 822-5109  
✉ [carenini@cs.ubc.ca](mailto:carenini@cs.ubc.ca)

### [Dr. Raymond T. Ng](#)

Professor  
Co-Director of Data Science Institute  
Dept. of Computer Science  
University of British Columbia  
2366 Main Mall, Vancouver, B.C.  
Canada V6T 1Z4  
☎ +1 (604) 822-2394  
✉ [rng@cs.ubc.ca](mailto:rng@cs.ubc.ca)

### [Dr. Jaideep Srivastava](#)

Professor  
University of Minnesota  
Keller 5-209  
☎ +1 (612) 625-4012  
✉ [srivasta@umn.edu](mailto:srivasta@umn.edu)