

# Hassan Sajjad | Scientist

## Publications

---

### Under-review Conference and Journal Papers

Firoj Alam, **Hassan Sajjad**, Muhammad Imran, and Ferda Ofli. "Standardizing and Benchmarking Crisis-related Social Media Datasets for Humanitarian Information Processing". In: *Underreview*. 2020.

Fahim Dalvi, **Hassan Sajjad**, Nadir Durrani, and Yonatan Belinkov. "Efficient Feature-based Transfer Learning". In: *Underreview*. 2020.

Prakhar Ganesh et al. "Compressing Large-Scale Transformer-Based Models: A Case Study on BERT". In: *Underreview*. ([link](#)). 2020.

**Hassan Sajjad**, Fahim Dalvi, Nadir Durrani, and Preslav Nakov. "Poor Man's BERT: Smaller and Faster Transformer Models". In: *Underreview*. 2020.

John Wu, Yonatan Belinkov, **Hassan Sajjad**, Nadir Durrani, Fahim Dalvi, and James Glass. "Similarity Analysis of Contextual Word Representation Models". In: *Underreview*. 2020.

### Peer-reviewed Journal Papers

Yonatan Belinkov, Nadir Durrani, Fahim Dalvi, **Hassan Sajjad**, and James Glass. "On the Linguistic Representational Power of Neural Machine Translation Models". In: *Computational Linguistics (CL)* (2020). ([link](#)).

Abdul Rafae, Asim Karim, **Hassan Sajjad**, Faisal Kamiran, and Jia Xu. "A Clustering Framework for Lexical Normalization of Roman Urdu". In: *Natural Language Engineering (NLE)* (2020).

Shafiq Joty, Nadir Durrani, **Hassan Sajjad**, and Ahmed Abdelali. "Domain Adaptation using Neural Network Joint Model". In: *Computer Speech and Language* 45.C (2017). ([link](#)). ISSN: 0885-2308.

**Hassan Sajjad**, Helmut Schmid, Alexander Fraser, and Hinrich Schütze. "Statistical models for unsupervised, semi-supervised and supervised transliteration mining". In: *Computational Linguistics* 43.2 (2017). ([link](#)).

Walid Magdy, **Hassan Sajjad**, Tarek El-Ganainy, and Fabrizio Sebastiani. "Bridging social media via distant supervision". In: *Social Network Analysis and Mining* 35.5 (2015). ([link](#)).

### Peer-reviewed Conference Papers

Nadir Durrani, Fahim Dalvi, **Hassan Sajjad**, Yonatan Belinkov, and Preslav Nakov. "One Size Does Not Fit All: Comparing NMT Representations of Different Granularities". In: *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. ([link](#)). Minneapolis, US, June 2019.

Hamdy Mubarak, Ahmed Abdelali, **Hassan Sajjad**, Younes Samih, and Kareem Darwish. "Highly Effective Arabic Diacritization using Sequence to Sequence Modeling". In: *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. ([link](#)). Minneapolis, US, June 2019.

D. Anthony Bau\*, Yonatan Belinkov\*, **Hassan Sajjad**, Fahim Dalvi, Nadir Durrani, and James Glass. "Identifying and Controlling Important Neurons in Neural Machine Translation". In: *International Conference on Learning Representations (ICLR)*. ([link](#)). New Orleans, US, May 2019.

Fahim Dalvi\*, Nadir Durrani\*, **Hassan Sajjad**\*, Yonatan Belinkov, D. Anthony Bau, and James Glass. "What is one Grain of Sand in the Desert? Analyzing Individual Neurons in Deep NLP Models". In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. ([link](#)). Honolulu, US, Mar. 2019.

Fahim Dalvi, Nadir Durrani, **Hassan Sajjad**, and Stephan Vogel. "Incremental Decoding and Training Methods for Simultaneous Translation in Neural Machine Translation". In: *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. ([link](#)). New Orleans, US, June 2018.

Yonatan Belinkov, Nadir Durrani, Fahim Dalvi, **Hassan Sajjad**, and James Glass. "What do Neural Machine Translation Models Learn about Morphology?" In: *Proceedings of the 55th Conference of the Association for Computational Linguistics (ACL)*. ([link](#)). Vancouver, Canada, Aug. 2017.

**Hassan Sajjad**, Fahim Dalvi, Nadir Durrani, Ahmed Abdelali, Yonatan Belinkov, and Stephan Vogel. "Challenging Language-Dependent Segmentation for Arabic: An Application to Machine Translation and Part-of-Speech Tagging". In: *Proceedings of the 55th Conference of the Association for Computational Linguistics (ACL)*. ([link](#)). Vancouver, Canada, Aug. 2017.

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 11th International AAAI Conference on Web and Social Media (ICWSM)*. ([link](#)). Montreal, Canada, May 2017.

Yonatan Belinkov, Lluís Màrquez, **Hassan Sajjad**, Nadir Durrani, Fahim Dalvi, and James Glass. "Evaluating Layers of Representation in Neural Machine Translation on Part-of-Speech and Semantic Tagging Tasks". In: *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP)*. ([link](#)). Taipei, Taiwan, Nov. 2017.

Fahim Dalvi, Nadir Durrani, **Hassan Sajjad**, Yonatan Belinkov, and Stephan Vogel. "Understanding and Improving Morphological Learning in the Neural Machine Translation Decoder". In: *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP)*. ([link](#)). Taipei, Taiwan, Nov. 2017.

**Hassan Sajjad** et al. "Eyes Don't Lie: Predicting Machine Translation Quality Using Eye Movement." In: *Proceedings of the 15th Annual Conference of the North American Chapter of the Association of Computational Linguistics: Human Language Technologies (NAACL-HLT)*. ([link](#)). San Diego, US, June 2016.

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)*. ([link](#)). Osaka, Japan, Dec. 2016.

**Hassan Sajjad**, Francisco Guzmán, and Stephan Vogel. "An Empirical Study: Post-editing Effort for English to Arabic Hybrid Machine Translation". In: *Proceedings of the Association for Machine Translation in the Americas (AMTA)*. ([link](#)). Austin, US, Oct. 2016.

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)*. ([link](#)). Lisbon, Portugal, Sept. 2015.

Abdul Rafae, Abdul Qayyum, Muhammad Moeen Uddin, Asim Karim, **Hassan Sajjad**, and Faisal Kamiran. "An Unsupervised Method for Discovering Lexical Variations in Roman Urdu Informal Text." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. ([link](#)). Lisbon, Portugal, Sept. 2015.

Walid Magdy, **Hassan Sajjad**, Tarek El-Ganainy, and Fabrizio Sebastiani. "Distant Supervision for Tweet Classification Using YouTube Labels". In: *Proceedings of the Ninth International AAAI Conference on Web and Social Media (ICWSM)*. ([link](#)). Oxford, UK, May 2015.

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 15th Machine Translation Summit (MT Summit XV)*. ([link](#)). Florida, USA, Nov. 2015.

Ahmed Abdelali, Francisco Guzman, **Hassan Sajjad**, and Stephan Vogel. "The AMARA Corpus: Building Parallel Language Resources for the Educational Domain". In: *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC)*. ([link](#)). Reykjavik, Iceland, May 2014.

Mohammad Moeen Uddin, Mohammad Imran, and **Hassan Sajjad**. "Understanding Types of Users on Twitter". In: *Proceedings of the 6th ASE International Conference in Social Computing (SocialCom)*. ([link](#)). Stanford, USA, May 2014.

Nadir Durrani, **Hassan Sajjad**, Hieu Hoang, and Philipp Koehn. "Integrating an Unsupervised Transliteration Model into Statistical Machine Translation". In: *Proceedings of the 15th Conference of the European Chapter of the ACL (EACL)*. ([link](#)). Gothenburg, Sweden, Apr. 2014.

Kareem Darwish, **Hassan Sajjad**, and Hamdy Mubarak. "Verifiably Effective Arabic Dialect Identification." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. ([link](#)). Doha, Qatar, Oct. 2014.

**Hassan Sajjad**, Kareem Darwish, and Yonatan Belinkov. "Translating Dialectal Arabic to English". In: *Proceedings of the 51st Conference of the Association for Computational Linguistics (ACL)*. ([link](#)). Sofia, Bulgaria, Aug. 2013.

**Hassan Sajjad**, Alexander Fraser, and Helmut Schmid. "A Statistical Model for Unsupervised and Semi-supervised Transliteration Mining". In: *Proceedings of the 50th Conference of the Association for Computational Linguistics (ACL)*. ([link](#)). Jeju, Korea, July 2012.

**Hassan Sajjad**, Patrick Pantel, and Michael Gamon. "Underspecified Query Refinement via Natural Language Question Generation". In: *Proceedings of the 24th International Conference on Computational Linguistics (COLING)*. ([link](#)). Mumbai, India, Dec. 2012.

**Hassan Sajjad**, Alexander Fraser, and Helmut Schmid. "An Algorithm for Unsupervised Transliteration Mining with an Application to Word Alignment". In: *Proceedings of the 49th Conference of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT)*. ([link](#)). Portland, OR, USA, June 2011.

**Hassan Sajjad**, Nadir Durrani, Helmut Schmid, and Alexander Fraser. "Comparing Two Techniques for Learning Transliteration Models Using a Parallel Corpus". In: *Proceedings of 5th International Joint Conference on Natural Language Processing (IJCNLP)*. ([link](#)). Chiang Mai, Thailand, Nov. 2011.

Nadir Durrani, **Hassan Sajjad**, Alexander Fraser, and Helmut Schmid. "Hindi-to-Urdu Machine Translation through Transliteration". In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*. ([link](#)). Uppsala, Sweden, July 2010.

**Hassan Sajjad** and Helmut Schmid. "Tagging Urdu Text with Parts of Speech: A Tagger Comparison". In: *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. ([link](#)). Athens, Greece, Apr. 2009.

## Peer-reviewed Demo Conference Papers

Hamdy Mubarak, Ahmed Abdelali, Kareem Darwish, Mohamed Eldesouki, Younes Samih, and **Hassan Sajjad**. "A System for Diacritizing Four Varieties of Arabic". In: *In Proceedings of the Empirical Methods in Natural Language Processing (EMNLP)*. ([link](#)). Nov. 2019.

Fahim Dalvi et al. "NeuroX: A Toolkit for Analyzing Individual Neurons in Neural Networks". In: *AAAI Conference on Artificial Intelligence (AAAI)*. ([link](#)). Honolulu, USA, Jan. 2019.

Fahim Dalvi et al. "QCRI Live Speech Translation System". In: *In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. ([link](#)). Apr. 2017.

Renars Liepins et al. "The SUMMA Platform Prototype". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. ([link](#)). Apr. 2017.

## Peer-reviewed Workshop Papers

**Hassan Sajjad**, Nadir Durrani, Fahim Dalvi, Yonatan Belinkov, and Stephan Vogel. "Neural Machine Translation Training in a Multi-Domain Scenario". In: *Proceedings of the 14th International Workshop on Spoken Language Translation (IWSLT)*. ([link](#)). Tokyo, Japan, Dec. 2017.

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: *Proceedings of the 4th International Workshop on Social Web for Disaster Management (SWDM)*. ([link](#)). Indianapolis, US, Oct. 2016.

Wajdi Zaghouani, Ahmed Abdelali, Francisco Guzmán, and **Hassan Sajjad**. "Normalizing Mathematical Expressions to Improve the Translation of Educational Content". In: *Proceedings of the AMTA 2016 Workshop Semitic Machine Translation (SeMaT)*. ([link](#)). Austin, US, Oct. 2016.

Francisco Guzmán, Ahmed Abdelali, Irina Temnikova, **Hassan Sajjad**, and Stephan Vogel. "How do Humans Evaluate Machine Translation". In: *Proceedings of the Workshop on Machine Translation (WMT)*. ([link](#)). Lisbon, Portugal, Sept. 2015.

Kamla Al-Mannai, **Hassan Sajjad**, Alaa Khader, Fahad Al Obaidli, Preslav Nakov, and Stephan Vogel. "Unsupervised word segmentation improves dialectal Arabic to English machine translation". In: *Proceedings of the Workshop of Arabic Natural Language Processing (ANLP)*. ([link](#)). Doha, Qatar, Oct. 2014.

Francisco Guzmán, **Hassan Sajjad**, Stephan Vogel, and Ahmed Abdelali. "The AMARA Corpus: Building Resources for Translating the Web's Educational Content". In: *Proceedings of the 10th International Workshop on Spoken Language Technology (IWSLT)*. ([link](#)). Heidelberg, Germany, Dec. 2013.

### Shared-task Workshop Papers

Xian Li et al. "Findings of the First Shared Task on Machine Translation Robustness". In: *Proceedings of the Fourth Conference on Machine Translation, Volume 2: Shared Task Papers*. ([link](#)). Florence, Italy, Aug. 2019.

Houda Bouamor and **Hassan Sajjad**. "H2@BUCC18: Parallel Sentence Extraction from Comparable Corpora Using Multilingual Sentence Embeddings". In: *Proceedings of the 11th Workshop on Building and Using Comparable Corpora (BUCC)*. ([link](#)). Miyazaki, Japan, May 2018.

Nadir Durrani, Fahim Dalvi, **Hassan Sajjad**, and Stephan Vogel. "QCRI's Machine Translation Systems for IWSLT'2016". In: *Proceedings of the 13th International Workshop on Spoken Language Translation (IWSLT)*. ([link](#)). Seattle, USA, Dec. 2016.

Mohamed Eldesouki, Fahim Dalvi, **Hassan Sajjad**, and Kareem Darwish. "QCRI @ DSL 2016: Spoken Arabic Dialect Identification Using Textual". In: *Proceedings of the 3rd Workshop on NLP for Similar Languages, Varieties and Dialects*. ([link](#)). Osaka, Japan, Dec. 2016.

Houda Bouamor, **Hassan Sajjad**, Nadir Durrani, and Kemal Oflazer. "QCMUQ@QALB-2015 Shared Task: Combining Character level MT and Error-tolerant Finite-State Recognition for Arabic Spelling Correction". In: *Proceedings of the Workshop of Arabic Natural Language Processing (ANLP)*. ([link](#)). Beijing, China, July 2015.

**Hassan Sajjad** et al. "QCN Egyptian Arabic to English Machine Translation System for NIST OpenMT15". In: *Workshop of NIST OpenMT15*. ([link](#)). Washington DC, US, June 2015.

Nadir Durrani, Helmut Schmid, Alexander Fraser, **Hassan Sajjad**, and Richárd Farkas. "Munich-Edinburgh-Stuttgart Submissions of OSM Systems at WMT13". In: *Proceedings of the Eighth Workshop on Statistical Machine Translation (WMT)*. ([link](#)). Sofia, Bulgaria, Aug. 2013.

**Hassan Sajjad**, Svetlana Smekalova, Nadir Durrani, Alexander Fraser, and Helmut Schmid. "QCRI-MES Submission at WMT13: Using Transliteration Mining to Improve Statistical Machine Translation". In: *Proceedings of the Eighth Workshop on Statistical Machine Translation (WMT)*. ([link](#)). Sofia, Bulgaria, Aug. 2013.

Marion Weller et al. "Munich-Edinburgh-Stuttgart Submissions at WMT13: Morphological and Syntactic Processing for SMT". In: *Proceedings of the Eighth Workshop on Statistical Machine Translation (WMT)*. ([link](#)). Sofia, Bulgaria, Aug. 2013.

**Hassan Sajjad** et al. "QCRI at IWSLT 2013: Experiments in Arabic-English and English-Arabic Spoken Language Translation". In: *Proceedings of the 10th International Workshop on Spoken Language Technology (IWSLT)*. ([link](#)). Heidelberg, Germany, Dec. 2013.