NADIR DURRANI

ARABIC LANGUAGE TECHNOLOGIES

Qatar Computing Research Institute
HBKU – Qatar Foundatation
+97450449634
ndurrani@hbku.edu.qa
Webpage
Google Scholar

EDUCATION

Doctor of Philosophy | ***Computer Science (Magna Cum Laude)*** April 2008 – Nov 2012
University of Stuttgart Stuttgart, Germany
Thesis Title: A Joint Translation Model with Integrated Reordering
IMS Best Thesis Award 2012 **GSCL Best Thesis Award 2014**
Committee: Professor Hinrich Schütze, Professor Alexander Fraser, Dr. Helmut Schmid, Professor François Yvon
Professor Dr. Bernhard Mitschang and Dr. Martin Fuchs

Masters of Science | ***Computer Science (Silver Medal)*** Aug 2005 – Aug 2007
National University of Computer and Emerging Sciences (FAST) Lahore, Pakistan
Thesis Title: Typology of word and automatic word segmentation in Urdu Supervisor: Professor Sarmad Hussain

Bachelors of Science | ***Computer Science*** Aug 2000 – Dec 2004
National University of Computer and Emerging Sciences (FAST) Lahore, Pakistan

RESEARCH INTERESTS

Explainable AI: Analyzing deep neural models with a focus on interpretation techniques like prompting, diagnostic classifiers, linguistic correlation analysis, unsupervised methods, clustering methods, ablation analysis, similarity analysis, safety steering, causation and behavioral studies, robustness and adversarial attacks, model editing

Machine Translation: Statistical and Neural MT with a focus on reordering, domain adaptation, transliteration, dialectal translation, pivoting, closely related and morphologically rich languages, low resource translation, spoken language translation and stream decoding

Other NLP: Word Segmentation (Urdu, Lao, Arabic), Speech Synthesis, Evaluation and Benchmarking of LLMs

Language Localization: localization of open source operating systems (desktop environments KDE/GNOME), word processing applications, chat tools, web browsers and web development tools, localization of domain names, word collation and font development

PROFESSIONAL EXPERIENCE

Senior Scientist April 2021 –
Qatar Computing Research Institute Doha, Qatar

- FANAR Shaheen – LLM-based machine translation
- LLMeBench – a framework for Benchmarking LLMs
- NeuroX – a toolkit for interpreting deep NLP models
- Shaheen – Dialect-to-English translation system
- NatiQ – Arabic Speech Synthesis System

Scientist Dec 2014 – Mar 2021
Qatar Computing Research Institute Doha, Qatar

- NeuroX – a toolkit for interpreting deep NLP models
- Shaheen, MSA–English

- QATS – Arabic-English Speech Translation System

Post-doctoral Research Associate Dec 2012 – Sept 2014
ILCC – University of Edinburgh Edinburgh, UK

- Integrated unsupervised transliteration mining and transliteration modeling into Moses
- Worked on integrating Ngram based translation and reordering into phrase-based Moses
- Worked on factored-based Ngram translation and reordering models
- Participated and won WMT and IWSLT Marathons consecutively in 2013-14

Research Intern Jan – April 2012
IBM – T.J. Watson Center New York, USA

- Worked on Egyptian Arabic to English Machine Translation

Research Staff April 2008 – Nov 2012
Institute of Natural Language Processing – University of Stuttgart Stuttgart, Germany

- Proposed a novel MT model called the Operation Sequence Model
- Wrote OSM decoder for Statistical MT and designed features for it
- Worked on Hindi-Urdu Machine Translation system
- Novel models for integrating transliteration into statistical MT decoder

Team Lead (Open Source Localization Project) Jan 2007 – Nov 2007
Center for Research in Urdu Language Processing – FAST Lahore, Pakistan

- Led a team of 10 employees in PAN Open Source Localization Project
- Localized Open Source Software (Open Office, GNome, KDE, SeaMonkey etc. for Urdu
- Deployed localized software to school kids in remote areas

Research Officer (PAN Localization Project) Aug 2005 – Dec 2006
Center for Research in Urdu Language Processing – FAST Lahore, Pakistan

- Worked on Urdu Domain Names
- Conducted a study on Collation of Languages from Developing Asia
- Conducted a survey of Language Computing in Asia

Project Coordinator Jan – Jun 2005
Science Technology and Environment Agency Vientiane, Laos

- Worked on Lao syllabification and collation algorithm
- Worked on Lao spelling correction
- Developed localized word pad for Lao

Research Assistant Aug 2003 – Dec 2004
Center for Research in Urdu Language Processing – FAST Lahore, Pakistan

- Worked on bootable Urdu Linux Distribution
- Developed fonts for Urdu
- Worked on Urdu phonetic keyboard

JOURNAL PUBLICATIONS

[J8] Shahad Al-Khalifa, **Nadir Durrani**, Hend Al-Khalifa, Firoj Alam (2025). The Landscape of Arabic Large Language Models (ALLMs): A New Era for Arabic Language Technology. *Communications of the ACM (CACM)* – Arab World Regional Special Section.

[J7] Shammur Absar Chowdhury, **Nadir Durrani**, Ahmed Ali (2024). Speech, Neuron-level analysis, Interpretability, Diagnostic classifier, AI explainability, End-to-end architecture. *Computer Speech and Language (CSL)* 83, Article 101539.

[J6] Nadir Durrani, Fahim Dalvi, Hassan Sajjad (2023). Discovering Salient Neurons in Deep NLP Models. *Journal of Machine Learning Research (JMLR)* 24:1–40, No. 362.

[J5] Hassan Sajjad, Fahim Dalvi, **Nadir Durrani**, Preslav Nakov (2023). On the Effect of Dropping Layers of Pre-trained Transformer Models. *Computer Speech and Language (CSL)* 77, Article 101429.

[J4] Hassan Sajjad*, **Nadir Durrani***, Fahim Dalvi* (2022). Neuron-level Interpretation of Deep NLP Models: A Survey. *Transactions of the Association for Computational Linguistics (TACL)* 10:1285–1303. **(*Equal Contribution—Order by draw)**

[J3] Yonatan Belinkov*, **Nadir Durrani***, Fahim Dalvi, Hassan Sajjad, Jim Glass (2020). On the Linguistic Representational Power of Neural Machine Translation Models. *Computational Linguistics (CL)* 46(1):1–57. **(*Equal Contribution—Alphabetic Order)**

[J2] Shafiq Joty, **Nadir Durrani**, Hassan Sajjad, Ahmed Abdelali (2017). Domain Adaptation Using Neural Joint Model. *Computer Speech and Language (CSL)* 45:161–179.

[J1] Nadir Durrani, Helmut Schmid, Alexander Fraser, Philipp Koehn (2015). The Operation Sequence Model – Combining N-gram-based and Phrase-based SMT. *Computational Linguistics (CL)* 41(2):157–186.

Conference Publications

[C47] Nadir Durrani*, Basel Mousi* and Fahim Dalvi (2025). Editing Across Languages: A survey of multilingual Knowledge Editing To appear in the *Proceedings of the 22nd Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China, November. **(*Equal Contribution—Alphabetic Order).**

[C46] Sabri Boughorbel, Fahim Dalvi, **Nadir Durrani** and Majd Hawasly (2025). Beyond the Leaderboard: Model Diffing for Understanding Performance Disparities in LLMs To appear in the *Proceedings of the 22nd Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China, November.

[C45] Basel Mousi, **Nadir Durrani** and Fahim Dalvi, (2025). An Exploration of Knowledge Editing for Arabic. To appear in the *Proceedings of the The Third Arabic Natural Language Processing Conference*, Suzhou, China, November

[C44] Asim Ersoy, Basel Mousi, Shammur Chowdhury, Firoj Alam, Fahim Dalvi, and **Nadir Durrani** (2025). From Words to Waves: Analyzing Concept Formation in Speech and Text-Based Foundation Models. In *Proceedings of the 26th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Rotterdam, Netherlands, August.

[C43] Basel Mousi, **Nadir Durrani**, Fatema Ahmad, Md. Arid Hasan, Maram Hasanain, Tameem Kabbani, Fahim Dalvi, Shammur Absar Chowdhury and Firoj Alam (2025). AraDiCE: Benchmarks for Dialectal and Cultural Capabilities in LLMs. In *Proceedings of the The 31st International Conference on Computational Linguistics (COLING)*, Abu Dhabi, UAE, January.

[C42] Xuemin Yu, Fahim Dalvi, **Nadir Durrani**, Marzia Nouri, Hassan Sajjad (2024). Latent Concept-based Explanation of NLP Models. In *Proceedings of the 21st Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Florida, Miami, November.

[C41] Basel Mousi, **Nadir Durrani**, Fahim Dalvi, Majd Hawasly, Ahmed Abdelali (2024). Exploring Alignment in Shared Cross-lingual Spaces. In *Proceedings of the 62th Annual Conference of the Association for Computational Linguistics (ACL)*, Bangkok, Thailand, August.

[C40] Majd Hawasly, Fahim Dalvi, **Nadir Durrani** (2024). Scaling up Discovery of Latent Concepts in Deep NLP Models. In *In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, St Julians, Malta, March.

[C39] Ahmed Abdelali*, Hamdy Mubarak*, Shammur Absar Chowdhury, Maram Hasanain, Basel Mousi, Sabri Boughorbel, Samir Abdaljalil, Yassine El Kheir, Daniel Izham, Fahim Dalvi, Majd Hawasly, Nizi Nazar, Yousseif Elshahawy, Ahmed Ali, **Nadir Durrani**, Natasa Milic-Frayling, Firoj Alam (2024). LAraBench: Benchmarking Arabic AI with Large Language Models. In *In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, St Julians, Malta, March **(*Equal Contribution—Alphabetic Order).**

[C38] Yimin Fan, Fahim Dalvi, **Nadir Durrani**, Hassan Sajjad (2023). Evaluating Neuron Interpretation Methods of NLP Models. In *Proceedings of the Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, December.

[C37] Basel Mousi, **Nadir Durrani**, Fahim Dalvi (2023). Can LLMs Facilitate Interpretation of Pre-trained Language Models? In *Proceedings of the 20th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Singapore, December.

[C36] Nadir Durrani, Hassan Sajjad, Fahim Dalvi, Firoj Alam (2022). On the Transformation of Latent Spaces in Fine-Tuned NLP Models. In *Proceedings of the 19th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, UAE, December.

[C35] Hassan Sajjad, Firoj Alam, Fahim Dalvi, **Nadir Durrani**, (2022). Effect of Post-processing on Contextualized Word Representations. In *Proceedings of the The 29th International Conference on Computational Linguistics (COLING)*. Gyeongju, South Korea, October

[C34] Hassan Sajjad, **Nadir Durrani**, Fahim Dalvi, Firoj Alam, Abdul Rafae Khan, Jia Xu (2022). Analyzing Encoded Concepts in Transformer Language Models. In *Proceedings of the 19th Annual Conference of the North American Chapter of the Association of Computational Linguistics: Human Language Technologies (NAACL)*. Seattle, USA, July

[C33] Fahim Dalvi*, Abdul Rafae Khan*, Firoj Alam, **Nadir Durrani**, Jia Xu, Hassan Sajjad (2022). Discovering Latent Concepts Learned in BERT. In *Proceedings of the 10th International Conference on Learning Representations (ICLR)*. Virtual, April **(*Equal Contribution—Alphabetic Order)**.

[C32] Firoj Alam, Shaden Shaar, Fahim Dalvi, Hassan Sajjad, Alex Nikolov, Hamdy Mubarak, Giovanni Da San Martino, Ahmed Abdelali, **Nadir Durrani**, Kareem Darwish, Abdulaziz Al-Homaid, Wajdi Zaghouani, Tommaso Caselli, Gijs Danoe, Friso Stolk, Britt Bruntink, Preslav Nakov (2021). Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society. In *Findings of the Empirical Methods in Natural Language Processing (EMNLP).* Virtual, November

[C31] Nadir Durrani, Hassan Sajjad, Fahim Dalvi (2021). How transfer learning impacts linguistic knowledge in deep NLP models. In *Findings of the Association for Computational Linguistics (ACL-IJCNLP).* Virtual, August

[C30] Firoj Alam, Fahim Dalvi, Shaden Shaar, **Nadir Durrani**, Hamdy Mubarak, Alex Nikolov, Giovanni Da San Martino, Ahmed Abdelali, Hassan Sajjad, Kareem Darwish, Preslav Nakov (2021). Fighting the COVID-19 Infodemic in Social Media: A Holistic Perspective and a Call to Arms. In *Proceedings of the 15th International Conference on Web and Social Media (ICWSM).* Virtual, August

[C29] Hassan Sajjad, Ahmed Abdelali, **Nadir Durrani**, Fahim Dalvi (2020). AraBench: Benchmarking Dialectal Arabic-English Machine Translation. In *Proceedings of the 28th International Conference on Computational Linguitics (COLING)*, Barcelona, Spain, December

[C28] Nadir Durrani, Hassan Sajjad, Fahim Dalvi, Yonatan Belinkov (2020). Analyzing Individual Neurons in Pre-trained Language Models. In *Proceedings of the 17th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Punta Cana, Dominican Republic, November.

[C27] Fahim Dalvi, Hassan Sajjad, **Nadir Durrani**, Yonatan Belinkov (2020). Analyzing Redundancy in Pretrained Transformer Models. In *Proceedings of the 17th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Punta Cana, Dominican Republic, November.

[C26] John M Wu*, Yonatan Belinkov*, Hassan Sajjad, **Nadir Durrani**, Fahim Dalvi and James Glass (2020). Similarity Analysis of Contextual Word Representation Models. In *Proceedings of the 58th Annual Conference of the Association for Computational Linguistics (ACL)*. Seattle, USA, July (*Equal Contribution—Alphabetic Order).

[C25] Nadir Durrani, Fahim Dalvi, Hassan Sajjad, Yonatan Belinkov, and Preslav Nakov (2019). One Size Does Not Fit All: Comparing NMT Representations of Different Granularities. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association of Computational Linguistics: Human Language Technologies (NAACL)*, Minneapolis, US, June

[C24] Anthony Bau*, Yonatan Belinkov*, Hassan Sajjad, Fahim Dalvi, **Nadir Durrani**, and James Glass (2019). Identifying and Controlling Important Neurons in Neural Machine Translation. In *Proceedings of the 7th International Conference on Learning Representations (ICLR)*. New Orleans, USA, May **(*Equal Contribution—Alphabetic Order)**.

[C23] Fahim Dalvi*, **Nadir Durrani***, Hassan Sajjad*, Yonatan Belinkov, D. Anthony Bau, and James Glass (2019). What is one Grain of Sand in the Desert? Analyzing Individual Neurons in Deep NLP Models. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. Honolulu, USA, Jan. **(*Equal Contribution—Alphabetic Order)**.

[C22] Fahim Dalvi*, **Nadir Durrani***, Hassan Sajjad, and Stephan Vogel (2018). 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 of Computational Linguistics: Human Language Technologies (NAACL)*, New Orleans, US, June **(*Equal Contribution—Alphabetic Order)**

[C21] Fahim Dalvi, **Nadir Durrani**, Hassan Sajjad, Yonatan Belinkov and Stephan Vogel (2017). Understanding and Improving Morphological Learning in the Neural Machine Translation Decoder. In *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP)*, Taipei, Taiwan, November.

[C20] Yonatan Belinkov, Lluís Màrquez, Hassan Sajjad, **Nadir Durrani**, Fahim Dalvi and James Glass (2017). Evaluating Layers of Representation in Neural Machine Translation on Part-of-Speech and Semantic Tagging Tasks. In *Proceedings of the 8th International Conference on Natural Language Processing (IJCNLP)*, Taipei, Taiwan, November

[C19] Yonatan Belinkov, **Nadir Durrani**, Fahim Dalvi, Hassan Sajjad, James Glass (2017). What do Neural Machine Translation Models Learn about Morphology? In Proceedings of the 55th Annual Conference of the Association for Computational Linguistics (ACL), Vancouver, Canada, July.

[C18] Hassan Sajjad, Fahim Dalvi, **Nadir Durrani**, Yonatan Belinkov, Ahmed Abdelali, Stephan Vogel (2017). Challenging Language-Dependent Segmentation for Arabic: An Application to Machine Translation and Part-of-Speech Tagging. In *Proceedings of the 55th Annual Conference of the Association for Computational Linguistics (ACL)*, Vancouver, Canada, July

[C17] Nadir Durrani, Hassan Sajjad, Shafiq Joty, Ahmed Abdelali (2016). A Deep Fusion Model for Domain Adaptation in Phrase-based MT. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, Osaka, Japan, December.

[C16] Hassan Sajjad, Francisco Guzmán, **Nadir Durrani**, Ahmed Abdelali, Houda Bouamor, Irina Temnikova and Stephan Vogel (2016). 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)*, San Diego, US, June

[C15] Ahmad Musleh, **Nadir Durrani**, Irina Temnikova, Preslav Nakov, Stephan Vogel and Osama Alsaad (2016). Enabling Medical Translation for Low-Resource Languages. In *Proceedings of the 16th Conference on Intelligent Text Processing and Computational Linguistics (CICLING)*, Konya, Turkey, April

[C14] Nadir Durrani, Hassan Sajjad, Shafiq Joty, Ahmed Abdelali, Stephan Vogel (2015). Using Joint Models for Domain Adaptation in Statistical Machine Translation In *Proceedings of the 15th MT Summit (AMTA)*, Florida, US, November.

[C13] Shafiq Joty, Hassan Sajjad, **Nadir Durrani**, Kamla Al-Mannai, Ahmed Abdelali, Stephan Vogel (2015). How to Avoid Unwanted Pregnancies: Domain Adaptation using Neural Network Models. In *Proceedings of the 12th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Lisbon, Portugal, September.

[C12] Nadir Durrani, Philipp Koehn, Helmut Schmid, Alexander Fraser (2014). Investigating the Usefulness of Generalized Word Representations in SMT Models. In *Proceedings of the 25th International Conference on Computational Linguistics (COLING)*, Dublin, Ireland, August.

[C11] Nadir Durrani and Philipp Koehn (2014). Improving Machine Translation through Triangulation and Transliteration. In *Proceedings of the 17th Annual Conference of the European Association for Machine Translation (EAMT)*, Dubrovnik, Croatia, June.

[C10] Nadir Durrani, Hassan Sajjad, Hieu Hoang, Philipp Koehn (2014). Integrating an Unsupervised Transliteration Model into Statistical Machine Translation. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Gothenburg, Sweden, April.

[C9] Nadir Durrani, Yaser Al-Onaizan and Abraham Ittycheriah (2014). Improving Egyptian-to-English SMT by mapping Egyptian into MSA. In *Proceedings of the 14th Conference on Intelligent Text Processing and Computational Linguistics (CICLING)*, Kathmandu, Nepal, April.

[C8] Nadir Durrani, Alexander Fraser, Helmut Schmid, Hieu Hoang, Philipp Koehn (2013). Can Markov Models Over Minimal Translation Units Help Phrase-Based SMT? In *Proceedings of the 51st Annual Conference of the Association for Computational Linguistics (ACL)*, Sofia, Bulgaria, August.

[C7] Nadir Durrani, Alexander Fraser, Helmut Schmid (2013). Model with Minimal Translation Units, But Decode With Phrases. In *Proceedings of the 14th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, Atlanta, Georgia, USA, June.

[C6] Hassan Sajjad, **Nadir Durrani**, Helmut Schmid, Alexander Fraser, (2011). Comparing Two Techniques for Learning Transliteration Models using a Noisy Parallel Corpus. To Appear, In *Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP)*, Chiang Mai, Thailand, November.

[C5] Nadir Durrani, Helmut Schmid, Alexander Fraser, (2011). A Joint Sequence Translation Model with Integrated Reordering. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL-HLT)*, Portland, Oregon, USA, June.

[C4] Nadir Durrani, Hassan Sajjad, Alexander Fraser, Helmut Schmid (2010). Hindi-to-Urdu Machine Translation Through Transliteration. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, Uppsala, Sweden, July.

[C3] Nadir Durrani and Sarmad Hussain (2010). Urdu Word Segmentation. In *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, Los Angeles, California, USA, June.

[C2] Sarmad Hussain and **Nadir Durrani** (2006). Urdu Domain Names, In *Proceedings of 10th IEEE Multitopic Conference 2006 (INMIC)*, Islamabad, Pakistan, December.

[C1] Nadir Durrani (2006). System of Grammatical Relations in Urdu, In *Proceedings 12th Himalayan Language Symposium & 27th Annual Conference of Linguistic Society of Nepal*, Kathmandu, Nepal, November.

WORKSHOP PUBLICATIONS

[W23] Ahmed Abdelali, **Nadir Durrani**, Fahim Dalvi, Hassan Sajjad (2022). Post-hoc analysis of Arabic transformer models. In *Proceedings of the Fifth Workshop on BlackboxNLP,* Abu Dhabi, UAE, December

[W22] Ahmed Abdelali, **Nadir Durrani**, Cenk Demiroglu, Fahim Dalvi, Hamdy Mubarak, Kareem Darwish (2022). NatiQ: An End-to-end Text-to-speech system for Arabic. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop* Abu Dhabi, UAE, December

[W21] Lucia Specia, Zhenhai Li, Juan Pino, Vishrav Chudhary, Francisco Guzman, Paul Michel, Graham Neubig, Hassan Sajjad, **Nadir Durrani**, Yonatan Belinkov, Philipp Koehn, and Xian Li (2020). Findings of the WMT 2020 Shared Task on Machine Translation Robustness. In *Proceedings of the Fifth Conference on Machine Translation, Volume 2: Shared Task Papers.* Punta Cana, Dominican Republic, November

[W20] Ebrahim Ansari, Amittai Axelrod, Nguyen Bach, Ondrej Bojar, Roldano Cattoni, Fahim Dalvi, **Nadir Durrani**, Marcello Federico, Christian Federmann, Jiatao Gu, Fei Huang, Kevin Knight, Xutai Ma, Ajay Nagesh, Matteo Negri, Jan Niehues, Juan Pino, Elizabeth Salesky, Xing Shi, Sebastian Stüker, Marco Turchi, Changhan Wang (2020). Findings of the IWSLT 2020 Evaluation Campaign. In *Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT 2020)*, Seattle, USA

[W19] Xian Li, Paul Michel, Antonios Anastasopoulos, Yonatan Belinkov, **Nadir Durrani**, Orhan Firat, Philipp Koehn, Graham Neubig, Juan Pino, and Hassan Sajjad (2019). Findings of the First Shared Task on Machine Translation Robustness. In *Proceedings of the Fourth Conference on Machine Translation, Volume 2: Shared Task Papers.* Florence, Italy, August

[W18] Fahim Dalvi*, **Nadir Durrani***, Hassan Sajjad*, Yonatan Belinkov, D. Anthony Bau, and James Glass (2019). What is one Grain of Sand in the Desert? Analyzing Individual Neurons in Deep NLP Models. In *Proceedings of the AAAI Workshop on Network Interpretability for Deep Learning.* Honolulu, USA, Jan. (*Equal Contribution—Alphabetic Order).

[W17] Anthony Bau*, Yonatan Belinkov*, Hassan Sajjad, Fahim Dalvi, **Nadir Durrani**, and James Glass (2018). Identifying and Controlling Important Neurons in Neural Machine Translation. In *Proceedings of the Interpretability and Robustness in Audio, Speech, and Language. NIPS Workshop (IRASL)*, Montreal, Canada, November (*Equal Contribution—Alphabetic Order).

[W16] **Nadir Durrani** and Fahim Dalvi (2017). Continuous Space Neural Reordering Models for Phrase-based MT. In *Proceedings of the 14th Workshop on Spoken Language Translation (IWSLT)*, Tokyo, Japan. December

[W15] Hassan Sajjad, **Nadir Durrani**, Fahim Dalvi, Yonatan Belinkov and Stephan Vogel (2017). Neural Machine Translation Training in a Multi-Domain Scenario. In *Proceedings of the 14th Workshop on Spoken Language Translation (IWSLT)*, Tokyo, Japan, December

[W14] **Nadir Durrani**, Fahim Dalvi, Hassan Sajjad, Stephan Vogel (2016). QCRI's Machine Translation Systems for IWSLT'2016. In *Proceedings of the 13th International Workshop on Spoken Language Translation (IWSLT)*, Seattle, USA, December

[W13] Houda Bouamor, Hassan Sajjad, **Nadir Durrani**, Kemal Oflazer (2015). 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)*, Beijing, China, July

[W12] Hassan Sajjad, **Nadir Durrani**, Francisco Guzman, Preslav Nakov, Ahmed Abdelali, Stephan Vogel, Wael Salloum, Ahmed El Kholy, Nizar Habash (2015). QCN Egyptian Arabic to English Machine Translation System for NIST OpenMT15, In the Workshop of NIST OpenMT15, Washington DC, USA, June

[W11] Alexandra Birch, Matthias Huck, **Nadir Durrani**, Nikolay Bogoychev, Philipp Koehn (2014). Edinburgh SLT and MT System Description for the IWSLT 2014 Evaluation. In *Proceedings of the 11th International Workshop on Spoken Language Translation (IWSLT)*, Lake Tahoe, USA, December

[W10] Markus Freitag, Joern Wuebker, Stephan Peitz, Hermann Ney, Matthias Huck, Alexandra Birch, **Nadir Durrani**, Philipp Koehn, Mohammed Mediani, Isabel Slawik, Jan Niehues, Eunah Cho, Alex Waibel, Nicola Bertoldi, Mauro Cettolo and Marcello Federico (2014). Combined Spoken Language Translation. In *Proceedings of the 11th International Workshop on Spoken Language Translation (IWSLT)*, Lake Tahoe, USA, December

[W9] **Nadir Durrani**, Barry Haddow, Kenneth Heafield, Philipp Koehn (2014). Edinburgh's Phrase-based Machine Translation Systems for WMT-14. In *Proceedings of the ACL 2014 Ninth Workshop on Statistical Machine Translation*, Baltimore, USA, June.

[W8] Markus Freitag, Stephan Peitz, Joern Wuebker, Hermann Ney, Matthias Huck, Rico Sennrich, **Nadir Durrani**, Maria Nadejde, Philip Williams, Philipp Koehn, Teresa Herrmann, Eunah Cho, Alex Waibel EU-BRIDGE MT: Combined Machine Translation. In *Proceedings of the ACL 2014 Ninth Workshop on Statistical Machine Translation*, Baltimore, USA, June.

[W7] Alexandra Birch, **Nadir Durrani**, Philipp Koehn (2013). Edinburgh SLT and MT System Description for the IWSLT 2013 Evaluation. In *Proceedings* of the Tenth International Workshop on Spoken Language Translation. Heidelberg, Germany, December.

[W6] Markus Freitag, Stephan Peitz, Joern Wuebker, Hermann Ney, **Nadir Durrani**, Matthias Huck, Philipp Koehn, Thanh-Le Ha, Jan Niehues, Mohammed Mediani, Teresa Herrman, Alex Waibel, Nicola Bertoldi, Mauro Cettolo and Marcello Federico (2013). EU-BRIDGE MT: Text Translation of Talks in EU-BRIDGE Project. In *Proceedings of the Tenth International Workshop on Spoken Language Translation*. Heidelberg, Germany, December.

[W5] Nadeem Khan, Waqas Anwar, Usama Ijaz Bajwa, **Nadir Durrani** (2013). English to Urdu Hierarchical Phrase-based Statistical Machine Translation. In *Proceedings of the Fourth Workshop on South and Southeast Asian Natural Language Processing*, Nagoya, Japan, October.

[W4] Nadir Durrani, Barry Haddow, Kenneth Heafield, Philipp Koehn (2013). Edinburgh's Machine Translation Systems for European Language Pairs. In *Proceedings of the ACL 2013 Eighth Workshop on Statistical Machine Translation*, Sofia, Bulgaria, August.

[W3] Nadir Durrani, Helmut Schmid, Alexander Fraser, Hassan Sajjad, Richárd Farkas (2013). Munich-Edinburgh-Stuttgart Submissions of OSM Systems at WMT13. In *Proceedings of the ACL 2013 Eighth Workshop on Statistical Machine Translation*, Sofia, Bulgaria, August.

[W2] Marion Weller, Max Kisselew, Svetlana Smekalova, Alexander Fraser, Helmut Schmid, **Nadir Durrani**, Hassan Sajjad and Richárd Farkas (2013). Munich-Edinburgh-Stuttgart Submissions at WMT13: Morphological and Syntactic Processing for SMT. In *Proceedings of the ACL 2013 Eighth Workshop on Statistical Machine Translation*, Sofia, Bulgaria, August.

[W1] Hassan Sajjad, Svetlana Smekalova, **Nadir Durrani**, Alexander Fraser, Helmut Schmid (2013). QCRI-MES Submission at WMT13: Using Transliteration Mining to Improve Statistical Machine Translation. In Proceedings of the ACL 2013 Eighth Workshop on Statistical Machine Translation, Sofia, Bulgaria, August.

DEMO PUBLICATIONS

[D9] Fahim Dalvi, Maram Hasanain, Sabri Boughorbel, Basel Mousi, Samir Abdaljalil, Nizi Nazar, Ahmed Abdelali, Shammur Absar Chowdhury, Hamdy Mubarak, Ahmed Ali, Majd Hawasly, **Nadir Durrani**, Firoj Alam (2024). LLMeBench: A Flexible Framework for Accelerating LLMs Benchmarking. In *In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, St Julians, Malta, March.

[D8] Fahim Dalvi, Hassan Sajjad, **Nadir Durrani** (2023). NeuroX Library for Neuron Analysis of Deep NLP Models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. Toronto, Canada, July

[D7] Fahim Dalvi, **Nadir Durrani**, Hassan Sajjad, Tamim Jaban, Mus'ab Husaini and Ummar Abbas (2023). NxPlain: A Web-based Tool for Discovery of Latent Concepts. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Dubrovnik, Croatia, May

[D6] Firoj Alam, Fahim Dalvi, **Nadir Durrani**, Hassan Sajjad, Abdul Rafae Khan, Jia Xu (2023). ConceptX: A Framework for Latent Concept Analysis. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*. Washington DC, USA, Feb

[D5] Fahim Dalvi, Avery Nortonsmith, D. Anthony Bau, Yonatan Belinkov, Hassan Sajjad, **Nadir Durrani**, James Glass (2019). NeuroX: A Toolkit for Analyzing Individual Neurons in Neural Networks. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. Honolulu, USA, Jan

[D4] Fahim Dalvi, Yifan Zhang, Sameer Khurana, **Nadir Durrani**, Hassan Sajjad, Ahmed Abdelali, Hamdy Mubarak, Ahmed Ali and Stephan Vogel (2017). QCRI Live Speech Translation System. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Valencia, Spain, April

[D3] Renars Liepins, Ulrich Germann, ... , **Nadir Durrani**,..., and Jeff Mitchell (2017). The SUMMA Platform Prototype. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Valencia, Spain, April

[D2] Ahmed Abdelali, **Nadir Durrani**, Francisco Guzmán (2016). iAppraise: A Manual Machine Translation Evaluation Environment. In *Proceedings of the 15th Annual Conference of the North American Chapter of the Association of*

Computational Linguistics: Human Language Technologies (NAACL), San Diego, US, June

[D1] Ahmed Abdelali, Kareem Darwish, **Nadir Durrani**, Hamdy Mubarak (2016). Farasa: A Fast and Furious Segmenter for Arabic. In *Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, San Diego, US, June

GRANTS AND PATENTS

- Principal Investigator, QNRF (with Darwish, Abdelali, Mubarak): TDF 03-0706-210009, Natiq: Arabic Text to Speech System, 6 months, USD 55000, September 2021
- **Patent (pending)**: Sajjad, Dalvi, Alam, Durrani, Khan, Xu, "*Latent Concept Analysis Method*", U.S. Patent Application No. US 18/132,224; filed Apr 7, 2023; published Oct 12, 2023.

PREPRINTS REPORTS (NON-PEER-REVIEWED)

[R12] FANAR-Team (2025). Fanar: An Arabic-Centric Multimodal Generative AI Platform, Qatar Computing Research Institute, Doha, Qatar, January. (Project Report)

[R11] Hassan Sajjad, Fahim Dalvi, **Nadir Durrani**, Preslav Nakov (2020). Poor Man's BERT: Smaller and Faster Transformer Models.

[R10] Nadeem Jadoon Khan, Waqas Anwar, **Nadir Durrani** (2013). Machine Translation Approaches and Survey for Indian Languages.

[R9] Nadir Durrani (2012). A Joint Translation Model with Integrated Reordering, University of Stuttgart, Institute of Natural Language Processing (IMS), Stuttgart, Germany, December. **(Doctoral Thesis – GSCL Best Doctoral Thesis, IMS Best Doctoral Thesis)**

[R8] Yaser Al-Onaizan, **Nadir Durrani**, Abe Ittycheriah (2012). Egyptian Arabic to MSA Normalization and Arabic Transliteration. Poster Bolt PI Meeting, Florida, USA, November.

[R7] Nadir Durrani (2007). Typology of Word and Automatic Word Segmentation in Urdu Text Corpus, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan, September. (Master's Thesis)

[R6] Sarmad Hussain and **Nadir Durrani** (2007). A Study on Collation of Languages from Developing Asia, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan.

[R5] Sarmad Hussain, **Nadir Durrani**, Sana Gul (2005). A Survey of Language Computing in Asia, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan.

[R4] Phonpasit Phissamay and **Nadir Durrani** (2005). Lao Collation, In Proceedings of PAN Localization Working Papers, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan.

[R3] Phonpasit Phissamay, Valaxay Dalolay, Chitaphone Chanhsililath, Oulaiphone Silimasak, Sarmad Hussain, **Nadir Durrani** (2005). Syllabification of Lao Script for Line Breaking, In Proceedings of PAN Localization Working Papers, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan.

[R2] Phonpasit Phissamay, **Nadir Durrani**, Thonglor Duangsavanh (2005). Lao Encoding Dilemma and Moving onto Unicode, In Proceedings of PAN Localization Working Papers, Center for Research in Urdu Language Processing, FAST-NUCES, Lahore, Pakistan.

[R1] Phonpasit Phissamay and **Nadir Durrani** (2005). Lao Character Set, In Proceedings of PAN Localization Working Papers, Center for Research in Urdu Language ProcessiDng, FAST-NUCES, Lahore, Pakistan.

Professional Activities

Senior Program Committee

- AAAI 2026: Senior Meta Reviewer (Area Chair)
- EMNLP 2025: Best Paper Award Committee
- IJCNLP 2025: Senior Area Chair in Interpretability and Analysis of Models for NLP
- EMNLP 2025: Senior Area Chair in Interpretability and Analysis of Models for NLP
- NAACL 2025: Senior Area Chair in Interpretability and Analysis of Models for NLP
- AAAI 2025: Senior Meta Reviewer (Area Chair)
- COLING 2025: Senior Area Chair in Interpretability and Analysis of Models for NLP
- Transactions of ACL (TACL) : Action Editor 2024
- EMNLP 2024: Senior Area Chair in Interpretability and Analysis of Models for NLP
    - Outstanding SAC Award
- AAAI 2024: Senior Meta Reviewer (Area Chair)
- NAACL 2024: Senior Area Chair in Interpretability and Analysis of Models for NLP
- COLING-LREC 2024: Area Chair in Applications Involving Language Resources and Evaluation
- EMNLP 2023: Senior Area Chair in Interpretability and Analysis of Models for NLP
- ACL 2023: Area Chair in Interpretability and Analysis of Models for NLP
- AAAI 2023: Meta Reviewer
- ACL 2022: ARR Action Editor
- AAAI 2022: Senior Meta Reviewer (Area Chair)
- EMNLP 2021: Area Chair in Natural Language Processing Applications
- ACL 2020: Area Chair in Machine Translation and Multilinguality
- AAAI 2020: Meta Reviewer
- CLT 2014: Publication Co-Chair

Workshop Co-organizing

- First Workshop on Machine Translation Robustness (@WMT 2019)

- Second Workshop on Machine Translation Robustness (@WMT 2020)
- First Workshop on Simultaneous Translation (@IWSLT 2020)

Tutorials

- Fine-grained Interpretation and Causation Analysis in Deep NLP Models(@NAACL 2021)

Journal Reviewing

- Computational Linguistics (CL): Standing Reviewer Since 2020
- Transactions of ACL (TACL) : Standing Reviewer Since 2020-2024
- Computer Speech and Language (CSL)
- Journal of Artificial Intelligence Research (JAIR)
- Transactions on Asian and Low-Resource Language Information Processing (TALLIP)
- Journal of Natural Language Engineering (JNLE)
- Language Resources and Evaluation (LREV)

Conference Reviewing

- Regular reviewer at ACL, NAACL, EACL, EMNLP and COLING since 2015
- Neural Information Processing Systems (NeurIPS 2020–2023, 2025)
- International Joint Conferences on Artificial Intelligence (IJCAI 2019–2020)
- International Conference on Representation Learning (ICLR 2020 – 2024)
- Conference on Artificial Intelligence (AAAI 2019)
- International Conference on Machine Learning (ICML 2021–2023)
- Conference on Computational Natural Language Learning (CoNLL 2018–2019))
- Conference of the European Association for Machine Translation (EAMT 2014–2017))
- Conference of the Association for Machine Translation (AMTA 2016–2018))
- Conference on Machine Translation (WMT 2016–2017)
- International Joint Conference on Natural Language Processing (IJCNLP 2017)
- Recent advances in natural language processing (RANLP 2017)
- Conference on Language and Technology (CLT 2014)

Workshop Reviewing

- Workshop on Statistical Machine Translation (WMT 2014–2015)
- Workshop on Multiword Expressions (MWE 2015)
- Workshop on Arabic Natural Language Processing (ANLP 2017)

Proposal Reviewing and Judging

- Annual Meeting of the Minds @ CMU 2024-25: Judge
- Qatar Debate Fellowship 2022
- Mauritius Research Council Proposal 2017
- Challenge 22 2015 (Sub))
- International Workshop on Semantic Evaluation (SemEval 2015)
- Nazarbayev University Research Council Proposal 2015

Contributions to Software

- **NeuroX:** Toolkit for interpreting deep NLP models
- **Shaheen:** Arabic – English and Dialect-to-English Neural Machine Translation System
- **FANAR:** An Arabic AI Large Language Model
- **LLMeBench:** A Flexible Framework for Accelerating LLMs Benchmarking
- **Moses:** Statistical Phrase-based Machine Translation System
- **OSM:** N-gram-based Statistical Machine Translation System
- **NatiQ:** Neural text-to-speech Arabic synthesizer
- **Farasa:** State-of-the-art Arabic segmenter

Honors and Awards

Outstanding Senior Area Chair Award EMNLP.	Miami 2024
King Salman Global Academy Prize for Arabic Language (Institution Track).	Riyadh 2023
NatiQ tech-transfer to Kinari AI	Doha 2022
Media coverage: pioneer dialectal Arabic translation Gulf Times, MENAFN, HBKU	Doha 2020
Shaheen tech-transfer to Kinari AI	Doha 2020
Shaheen deployed to BBC and Deutsche Welle as part of the SUMMA project	Doha 2019
Media coverage: Analyzing and controlling deep models	MIT News 2019
Best innovation award for Arabic–English Speech Translation at ARC	Doha 2017
Media coverage: Work on analyzing representations in deep models	MIT 2017
Best Paper Award at Open NIST 2015	Open NIST 2015
GSCL Best Computational Linguistics Dissertation Award of 1000 Euros	Kovens 2014
IMS best doctoral thesis award of 100 Euros	Stuttgart 2012
Got a 4 month Internship at IBM T.J. Watson Center	New York 2012
Funded by NAACL to present my paper "Urdu Word Segmentation" at the main conference	Los Angeles 2010
Secured a merit-based Scholarship for Doctoral Studies in Germany for a period of 4 years	Islamabad 2007
Graduated with Silver Medal (2nd Position in the batch)	Lahore 2007
Represented University in Software Exhibitions COMPAC	Islamabad 2004
Represented University in Software Exhibitions SOFTEC	Lahore 2003
Engineering Council Award for securing 2nd position at Higher Secondary Exam	Riyadh 2000
Awarded a prize of 200 Saudi Riyals by Pakistan International School Riyadh	Riyadh 2000
Earned a merit-based scholarship for intermediate studies	Riyadh 1998

Participation in Machine Translation Marathons

Ranked 1st for Arabic–English (both directions)	IWSLT 2016
Ranked 2nd for Egyptian–English (Chat, SMS and Speech Track)	Open NIST 2015
Ranked 1st for English–German and 3rd for English–French	IWSLT 2014
Ranked 1st for French–English	WMT 2014
Ranked 2nd for English–French, Czech–English, English–Czech	
Best constrained systems for CS–EN, RU–EN, EN–RU FR–EN, EN–FR, EN–HI	
Ranked 2nd for English–French	IWSLT 2013
Ranked 1st for Czech–English, French–English Spanish–English, English–French	WMT 2013
Ranked 2nd for French–English	
Best constrained systems for FR-EN, EN-FR, ES-EN, EN-ES, CS-EN, EN-CS	

Talks

NYU: Analyzing Contextualized Representations and Individual Neurons in Deep NLP	Virtual, 2020
Invited Speaker: 7th International Conference on Language and Technology	Lahore, Pakistan 2020
Invited Speaker: 6th International Conference on Language and Technology	Lahore, Pakistan 2016
IIT Bombay: The Operation Sequence Model	Mumbai, India 2015
IIT Bombay: Unsupervised Transliteration Model with Application to SMT	Mumbai, India 2015
University of Engineering and Technology: Statistical Machine Translation	Lahore, Pakistan 2014
SYSTRAN: A joint translation model with integrated reordering	Paris, France 2014
SDL: A joint translation model with integrated reordering	Los Angeles, USA 2014
WMT-14 Panel: The Operation Sequence Model	Baltimore, USA 2014

Trainings and Workshops

2nd Lisbon Machine Learning School	Lisbon 2012
Open Source Localization Training	Phenom Penh 2007
Asian Applied NLP for Linguistic Diversity and Language Resource (ADD)	Bangkok 2006
Summer School in Asian Language Processing	Lahore 2006
2nd Regional Training in PAN Localization	Siem Reap 2005
Workshop on Lao Localization	Vientiane 2005

Languages

- **Urdu** Native
- **English** Fluent
- **Punjabi** Fluent
- **Hindi** Fluent
- **German** Level: A–2