Sami Ul Haq

+353 899750234 | sami.haq2@mail.dcu.ie | https://sami-haq99.github.io

The Crescent, Park West, Dublin 12,(D12 TD80).

EDUCATION

Dublin City University (DCU)　　　　　　　　　　　　　　　　　　　Dublin, Ireland
PhD in Multi-modal Machine Translation, Advisor. Sheila Castilho　　　　　*Sep. 2023 – Aug 2027*

National University of Sciences and Technology (NUST)　　　　Islamabad, Pakistan
MS in Software Engineering　　　　　　　　　　　　　　　　　　　*Sep. 2014 – May 2017*

Quaid i Azam University (QAU)　　　　　　　　　　　　　　　Islamabad, Pakistan
MSC in Information Technology　　　　　　　　　　　　　　　　　*Jan. 2011 – Dec 2013*

EXPERIENCE

PhD Researcher　　　　　　　　　　　　　　　　　　　　Sept. 2023 – Aug. 2027
ADAPT Centre　　　　　　　　　　　　　　　　　　　　　　　*Dublin, Ireland*

- My research focuses on understanding Multi-modal Machine Translation (MMT) systems, and develop evaluation metrics for them.
- I aim to design an evaluation strategy for making MMT systems more interpretable, trustworthy and beneficial for society.

Tutoring and Lab Demonstration Roles　　　　　　　　　　Sept. 2024 – Present
Dublin City University　　　　　　　　　　　　　　　　　　　　*Dublin, Ireland*

- Translation Technology - **TRA1009:**　In current fall semester, I am working in the School of Applied Language and Intercultural Studies (SALIS) as Tutor, guiding Master's students in the use of various Computer-Assisted Translation (CAT) tools, including WordFast, Trados, and MemoQ. The labs also cover demonstration of human and automatic evaluation techniques for MT quality assessment.

Research Assistant　　　　　　　　　　　　　　　　　　Oct. 2018 – Sept. 2023
FJWU　　　　　　　　　　　　　　　　　　　　　　　　　*Rawalpindi, Pakistan*

- Explored and developed methods for context-aware Neural Machine Translation
- Evaluated and addressed challenges to enhance NLP solutions for low-resource languages, applied Transfer Learning, Multilingual Transfer, and Language Modeling
- Created interactive web interfaces for Machine Learning and MT applications using Django and Streamlit

Software Engineer　　　　　　　　　　　　　　　　　　　Jul. 2013 – Sep. 2018
Project Vision　　　　　　　　　　　　　　　　　　　　　　*Islamabad, Pakistan*

- Involved in a phased migration plan for Command & Control modules in an Air Space Management System, transitioning from Apple i386 to Intel-Xeon based architecture
- Improved existing back-bone encoding algorithm for efficient, reliable, and real-time data transmission of primary and secondary Radars
- Collaborated with a technical team on system integration with Airborne Warning and Control Systems (AWACS) as part of an international partnership.
- Assisted HR team in conducting first- and second-round interviews for software developer candidates

RESEARCH PROJECTS

Document Level Neural Machine Translation of Low Resourced Languages

- Project focused on utilizing techniques like Back-translation, Domain Adaptation to improve performance of NMT systems for low-resourced languages (Hindi-Marathi).
- **Publication**: "Document Level NMT of Low-Resource Languages with Backtranslation." Proceedings of the Fifth Conference on MT. 2020

Reverse Engineering UML Structural & Behavioral Features from Java Source Code

- Primary project focused on generation of Text-to-Model (T2M) transformation for static and dynamic models extraction. Object oriented constructs were transformed into EMF-UML Activity, Use-Case, Class and, State machine diagram.

- **Publication:** "A Model Driven Reverse Engineering Framework for Generating High Level UML Models from Java Source Code", IEEE Access, 2019, Volume 7, Pages 158931 - 158950

Model Driven Approach for Modeling & Analysis of Physical Infrastructure Protection
- The project involved building Domain Specific Modeling Language for physical security by extending industry standard UML. Annotated UML models then transformed into Bayesian Networks for Quantitative analysis of security configurations.
- **Publication:** "A Novel Approach for Modeling Security Aspects of Physical Infrastructures", International Conference on High-Performance Compilation, Computing and Communications (HP3C-2017)

PUBLICATIONS

[1] Sami Ul Haq, Rudali Huidrom, and Sheila Castilho. "DCU ADAPT at WMT24: English to Low-resource Multi-Modal Translation Task". In: ().

[2] Sami Ul Haq et al. "Context-Aware Neural Machine Translation using Selected Context". In: *2022 19th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*. IEEE. 2022, pp. 349–352.

[3] Sumbal Naz, Sadaf Abdul Rauf, and Sami Ul Haq. "FJWU Participation for the WMT21 Biomedical Translation Task". In: *Proceedings of the Sixth Conference on Machine Translation*. Online: Association for Computational Linguistics, Nov. 2021, pp. 857–862. URL: https://aclanthology.org/2021.wmt-1.86.

[4] Sami Ul Haq et al. "Document Level NMT of Low-Resource Languages with Backtranslation". In: *Proceedings of the Fifth Conference on Machine Translation*. Online: Association for Computational Linguistics, Nov. 2020, pp. 442–446. URL: https://aclanthology.org/2020.wmt-1.53.

[5] Sami Ul Haq et al. "Improving Document-Level Neural Machine Translation with Domain Adaptation". In: *Proceedings of the Fourth Workshop on Neural Generation and Translation*. Online: Association for Computational Linguistics, July 2020, pp. 225–231. DOI: 10.18653/v1/2020.ngt-1.27. URL: https://aclanthology.org/2020.ngt-1.27.

[6] Umair Sabir et al. "A model driven reverse engineering framework for generating high level UML models from java source code". In: *IEEE access* 7 (2019), pp. 158931–158950. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8890645.

ACADEMIC SERVICES

University services:
- **Co-PI** RAC (Research Academic Collaboration) Fatima Jinnah Women University, 2019-2021
- **FYP Supervisor** (Masters Thesis) AI Based Data Leak Prevention Systems, QAU, 2018-2019
- **Research Talk** (under-grad students) Fatima Jinnah Women University, 2020

Conferences:
- **Reviewer** IBCAST (2023), IBCAST(2022), IPCITI (2024)
- **Sub-reviewer** EAMT (2024), ACL (2023)

HONORS

PhD Scholarship Award : Science Foundation Ireland, D-real (SFI d-real), Ireland (2023-2027)
Merit Award (3MT) : College of E&ME, (NUST), Pakistan (2022)
Fully funded Summer College : Selected for 1-week International Summer College (INSC), Pakistan (2022)
Certificate of Appreciation : Project Vision, Pakistan (2017)
Vice Chancellor's Gold Medal : Institute of Information Technology, QAU, Pakistan (2016)
Academic Scholarship Award : Institute of Information Technology, QAU, Pakistan (2011-2013)

TECHNICAL SKILLS

Programming: Proficient in C, C++, Matlab, Java, Python and LATEX
Toolkits and Frameworks: Git, PyTorch,TensorFlow, fairseq, transformers, POSIX Threads, ElasticSearch
Developer ToolsVS Code, Visual Studio, IBM ALM , Eclipse