

# Sami Ul Haq

+353 899750234 | [sami.haq@adaptcentre.ie](mailto:sami.haq@adaptcentre.ie) | <https://sami-haq99.github.io>

Dublin 6, Dublin, Ireland.

## EDUCATION

---

### Dublin City University (DCU)

*PhD in Multi-modal Machine Translation, Advisor. Sheila Castilho*

Dublin, Ireland

*Sep. 2023 – Aug 2027*

### National University of Sciences and Technology (NUST)

*MS in Software Engineering*

Islamabad, Pakistan

*Sep. 2014 – May 2017*

## EXPERIENCE

---

### PhD Researcher

*ADAPT Centre*

Sept. 2023 – Aug. 2027

*Dublin, Ireland*

- My research focuses on understanding Multimodal 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

*Dublin City University*

Sept. 2024 – May. 2025

*Dublin, Ireland*

- Translation Technology - **TRA1009**: 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 covered demonstration of human and automatic evaluation techniques for MT quality assessment.

### Research Assistant

*FJ Women University*

Oct. 2018 – Sept. 2023

*Rawalpindi, Pakistan*

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

## PHD RESEARCH

---

### Multimodal Automatic Evaluation of MT

- \* I am currently developing a multimodal automatic evaluation metric to capture visual awareness and visual sense disambiguation capabilities of NLG systems, such as multimodal Machine Translation. The metric leverages vision-language models such as **Jina-CLIP** and **Aya-Vision** to extract representations from multiple modalities (text and images), enabling the computation of visually-informed semantic alignment between candidate and reference text. [Poster](#)

### Multimodal Human Evaluation of Machine Translation Quality

- \* During first year of my PhD, I designed and implemented a machine translation evaluation environment on **Amazon Mechanical Turk**, integrating **speech synthesis** for multimodal assessment. Analyzed over 30K+ segment ratings, demonstrating that the multimodal setup enhanced the detection of significant differences between MT systems.
- \* **Publication**: "Crowd-Sourced Evaluation of Machine Translation Quality." [Poster](#)

### Multimodal Translation Evaluation (MuTE) dataset

- \* Developed a cross-modal retrieval pipeline using VLMs and FAISS indexing to support high-speed semantic search across multilingual image-text pairs. Validated system performance through rigorous human-in-the-loop annotation, resulting in a highly semantically accurate benchmark dataset for model evaluation.
- \* **Publication**: "Vision-Grounded Evaluation of MMT using Vision Language Models." - Under-review - [Poster](#)

## PUBLICATIONS

---

- [1] Sami Haq, Sheila Castilho, and Yvette Graham. “Audio-Based Crowd-Sourced Evaluation of Machine Translation Quality”. In: *Proceedings of the Tenth Conference on Machine Translation*. Suzhou, China: Association for Computational Linguistics, Nov. 2025, pp. 52–63. ISBN: 979-8-89176-341-8. DOI: [10.18653/v1/2025.wmt-1.3](https://aclanthology.org/2025.wmt-1.3/). URL: <https://aclanthology.org/2025.wmt-1.3/>.
- [2] Sami Haq et al. “Long-context Reference-based MT Quality Estimation”. In: *Proceedings of the Tenth Conference on Machine Translation*. Suzhou, China: Association for Computational Linguistics, Nov. 2025, pp. 905–912. ISBN: 979-8-89176-341-8. DOI: [10.18653/v1/2025.wmt-1.64](https://aclanthology.org/2025.wmt-1.64/). URL: <https://aclanthology.org/2025.wmt-1.64/>.
- [3] Sami Ul Haq, Rudali Huidrom, and Sheila Castilho. “DCU ADAPT at WMT24: English to Low-resource Multi-Modal Translation Task”. In: *Proceedings of the Ninth Conference on Machine Translation*. Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 810–814. DOI: [10.18653/v1/2024.wmt-1.75](https://aclanthology.org/2024.wmt-1.75/). URL: [https://aclanthology.org/2024.wmt-1.75](https://aclanthology.org/2024.wmt-1.75/).
- [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](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](https://aclanthology.org/2020.ngt-1.27/). URL: [https://aclanthology.org/2020.ngt-1.27](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 & VOLUNTARY ROLES

---

### 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** (PhD students) School of Applied Language and Intercultural Studies (SALIS), DCU, 2025

### Conferences:

- \* **Co-organiser** StyGenAi-EAMT(2026)
- \* **Reviewer** EMNLP (2025), EMNLP(2024), IPCITI (2024), IBCAST (2023)
- \* **Sub-reviewer** COLM(2026), EAMT (2024), ACL (2023)

### Community Involvement:

- \* **ALONE** Providing weekly visitation support and companionship to an older person, 2025 – Present
- \* **Muslim Sisters of Éire** Distributed food and hygiene kits to people in need across Dublin

## HONORS

---

- Best Poster Award** : Scientific Conference, ADAPT ([ASC-2025](#)), Ireland (2025)
- 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:** Docker, AWS, Mturk, Appraise, Git, PyTorch, TensorFlow, fairseq, transformers, POSIX Threads, ElasticSearch, Meta FAISS
- Development Tools:** VS Code, Visual Studio, IBM ALM