

## Multi-modal-dialectal-cultural Arabic Understanding

### Project Description:

This project focuses on advancing technologies and resources for Arabic dialects and cultures across diverse modalities. By developing cutting-edge methods, it aims to bridge gaps in linguistic, cultural, and dialectal understanding for various applications, including training and evaluation of large language models (LLMs), education, anthropology, and beyond. We have an active team and we are seeking individuals with **any of the following skills and interests** to contribute to our multidisciplinary project:

- Machine learning and AI, particularly large language models (LLMs)
- Frontend or backend development
- Linguistics and language research

You don't need to have expertise in all these areas—if you're passionate about any one of these fields, we'd love to hear from you! Fluency in an Arabic dialect is a plus, but not required!

**Project Type:** Research and Engineering

### Internship Batch:

- **Batch 1:** May 11 to July 10, suitable for Education City students, i.e., CMUQ, TAMUQ and HBKU students
- **Batch 2:** May 25 to July 24, suitable for QU university students

### Duties/Activities:

We have different roles depending on the applicants interest:

1. NLP and Machine Learning
2. Backend development
3. Front-end/mobile development

### Required Skills:

Python, Basic deep learning/ML skills or Frontend/backend

**Preferred Intern Academic Level:**

Undergraduate or graduate

**Learning Opportunities:**

LLM training best practices

**Expected Team Size:**

Up to 10 interns, as each team would work on different capabilities

**Mentors**

Name: Ehsaneddin Asgari      email: [easgari@hbku.edu.qa](mailto:easgari@hbku.edu.qa)