



## Evaluating Cultural Relevance of Arabic AI Benchmarks Assessing Alignment with Arabic and Islamic Contexts

### Project Description

*Many AI benchmarks used for evaluating Arabic language models are direct translations of English benchmarks. While translation increases resource availability, it may introduce cultural bias, reduce contextual relevance, and overlook important aspects of Arabic and Islamic culture. This project aims to systematically analyze existing Arabic NLP benchmarks to assess their cultural appropriateness, contextual alignment, and representational fairness.*

*The internship will focus on identifying cultural gaps, imported biases, missing themes, and potential misalignment with Arabic and Islamic values, while proposing structured improvements and culturally grounded benchmark alternatives.*

### Project Type

*Research and Development (Arabic NLP + AI Evaluation + Cultural Analysis)*

### Objectives

- *Identify major Arabic NLP benchmarks translated from English.*
- *Analyze cultural alignment and contextual appropriateness of benchmark content.*
- *Detect implicit cultural bias or irrelevant scenarios.*
- *Evaluate representation of Arabic societies and Islamic contexts.*
- *Propose culturally grounded improvements and new benchmark dimensions.*

### Methodology

- *Literature review on benchmark design and cultural bias in AI evaluation.*
- *Systematic content analysis of selected Arabic benchmarks.*
- *Annotation framework for cultural relevance and sensitivity.*
- *Quantitative and qualitative analysis of identified issues.*
- *Development of recommendations and prototype culturally-aware benchmark samples.*

### Deliverables

- *Comparative analysis report of selected Arabic benchmarks.*

- *Cultural relevance assessment framework.*
- *List of identified biases and contextual gaps.*
- *Recommendations for benchmark redesign or extension.*
- *Prototype culturally-adapted benchmark samples.*

### **Required Skills (Any of the Following)**

- *Background in NLP, AI, or computational linguistics.*
- *Strong understanding of Arabic language and culture.*
- *Interest in AI ethics and bias analysis.*
- *Experience in qualitative or content analysis (preferred).*

### **Preferred Academic Level**

*Open to senior undergraduate students, MSc students, and fresh graduates in Computer Science, Artificial Intelligence, Linguistics, Islamic Studies, or related interdisciplinary fields.*

### **Learning Opportunities**

*Interns will gain hands-on experience in AI evaluation, bias analysis, and culturally-aware benchmark design. The project bridges AI research with cultural and ethical considerations, offering a unique interdisciplinary portfolio opportunity.*

### **Expected Team Size**

*2–4 interns*

### **Mentors:**

**Hamdy Mubarak** ([hmubarak@hbku.edu.qa](mailto:hmubarak@hbku.edu.qa)), Principal Software Engineer

**Majd Hawasly** ([mhawasly@hbku.edu.qa](mailto:mhawasly@hbku.edu.qa)), Scientist