Summer Internship Project Documentation - IKNEX Lab

26/08/2023

Team members:

Aiman Karim

Nibtahil Nafees

Overview

During the internship, the main goal was to convert uthmani Quranic scripts to those of Indo-Pak Quranic scripts. Due to the language being Arabic and having to deal with the Quran, special attention had to be given. The Conversions of diacritics was an ever present issue due to python's limitation. Research was conducted using encodings, unicodes and libraries and finally, the utmani text was successfully converted.

Goals

1. Convert the texts
2. Use encodings-unicodes for trickier diacritics
3. Use PyArabic library to convert the remaining diacritics

Difficulties

The first difficulty in conversion we felt was the changing of dagger alif. It had to be put on Indo-Pak script. The methods used to convert were mappings, encodings as well as logic use. Finally, it was changed by using PyArabic however, dagger alif on its own still remains unchanged.

The second we experienced was the changing of hamza above and hamza below diacritics. Neither encodings nor mappings were working. Using PyArabic helped convert the diacritics.

Some methodologies used

The first methods thought of to help were encodings that helped convert major diacritics like shadda. However, many others weren't being changed. A dictionary was created, one for Arabic alphabets and one for Arabic diacritics in XML file. This was done as we believed at that point, Python might have some limitations hence we thought to not involve python when diacritics were used.

End Result

After using the PyArabic library, most of our remaining issues were resolved. The only remaining issue is that of dagger alif when combined with other diacritics. The uthmani script converts successfully, a manual authentication was conducted. Work on authenticating via code had started.

The project proved to be a huge learning curve for the both of us, each of us exploring a new part of the coding world. Spending quality time with the mentor, where we explored the project in detail helped us make significant progress.