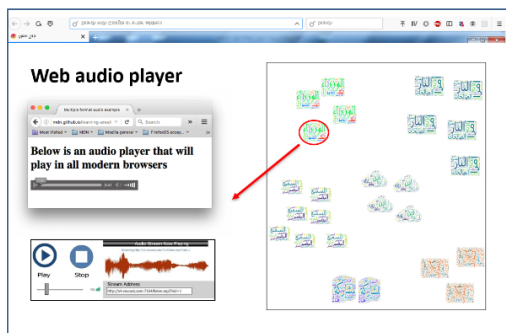


Linking visual summary of English/Arabic docs with text, audio, and video media

Project Description: Visualization plays an important role in advancing the analysis of Arabic dialect in written, audio or video formats. We develop an interface to support the categorization of these data by an analyst. Dialect data are visualized as small images on a map that the analyst can drag and drop to create the categories. We want to help the analyst getting a detailed view of the data when hovering on their small image. We need to link the snippet images back to their text, audio, image or video sources. We also need to generate images (wordcloud) from non-image data like text documents based on Natural Language Processing tools.



Duties/Activities: Implementing **frontend and backend Python and Javascript** code to support opening and display of **different format of sources (pdf, jpg, mp3, mp4...)** when hovering or clicking on small snippet images, using existing libraries. **Summarizing Arabic and English text** as Wordcloud images,

Required Skills: base knowledge of Javascript (D3.js) and Python 3 (Flask)

Preferred Intern Academic Level: BSc 3rd year or BSc 4rd year

Learning Opportunities: Advanced visualization and web frontend techniques; Natural Language Processing techniques

Expected Team Size: 2/3

Mentors Ala Abuthawabeh: aabuthawabeh@hbku.edu.qa and Michaël Aupetit: maupetit@hbku.edu.qa