UIP Centralized Database

Eng Kay Tze, Kong Mun Jun
Bachelor of Computer Science(Honours) in Data Science
Supervisor: Dr Lim Siew Mooi

Abstract

Academics focus more on basic research, education, and training, while industry focuses more on applied research and development in commercially viable application domains. In the field of AI in recent years, however, this distinction between the roles of academia and industry has eroded. Thus, it is aimed to form a centralized database with user interface that able to connect everyone in academia sector.

Problem Statement

1. Some university and ministry data are not real time information, not standardized and even fragmented
2. Researchers hardly find the universities and ministry information for their researches and projects.
3. Government need to have a strong connection with the ministry and universities to ensure the enhanced learning scope and contents.

Objectives

1. To crawl various valid information from the university, government agency into csv format.
2. To create a system that acts as a centralised database to gather pieces of universities and government information.
3. To provide convenient, efficient and accurate visualisation results to the third party.

Design and Methodology

Evolutionary Prototype Model are used in this project, UIP Centralized Database, as this model allow this project to be continuously build and refining till the best condition of the system are developed.

Hardware and Software requirements are quite common in this project. Basic desktop, mouse and keyboard are only required for hardware while python 3.8 and html5 will be used as Software which also easily to be installed.

Target Market for this project are listed below:

1. Universities researchers
2. Academic community
3. Centres of Excellence of Universities
4. Government agencies and Ministries
5. SMEs and Corporation in 15 industry Groups
6. Investors
7. Undergraduates and Postgraduate students

Construction and Testing

Visual Studio Code and Google Colab are used to construct the whole project to process the development of data cleaning and user-interface. Google colab mainly used for data sorting and data cleaning with python3.8 while Visual Studio Code is used to construct the database user-interface with the support of HTML5, CSS and JavaScript.

BeautifulSoup4 and Selinium also used to assist in Data Scraping module by scraping from any website and store the data into another tools known as Microsoft Excel with csv format.

User-Interface were created in a simplistic style by showing the features of the module obviously. Information in the module also went through every process of beautification before storing into database. White Box testing and Black Box testing also used to refine the module in the website to avoid bugs being published.

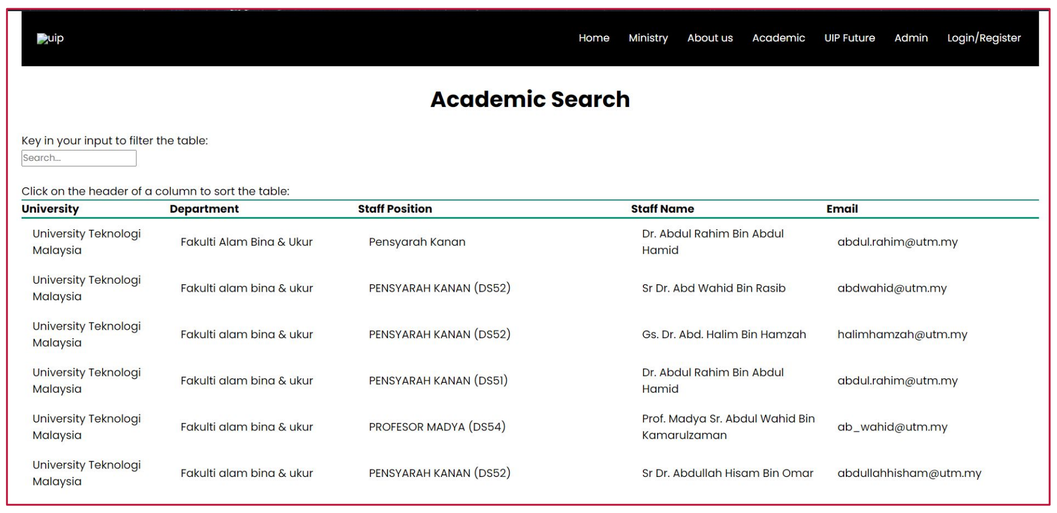

Figure 1: Academic Module Webpage

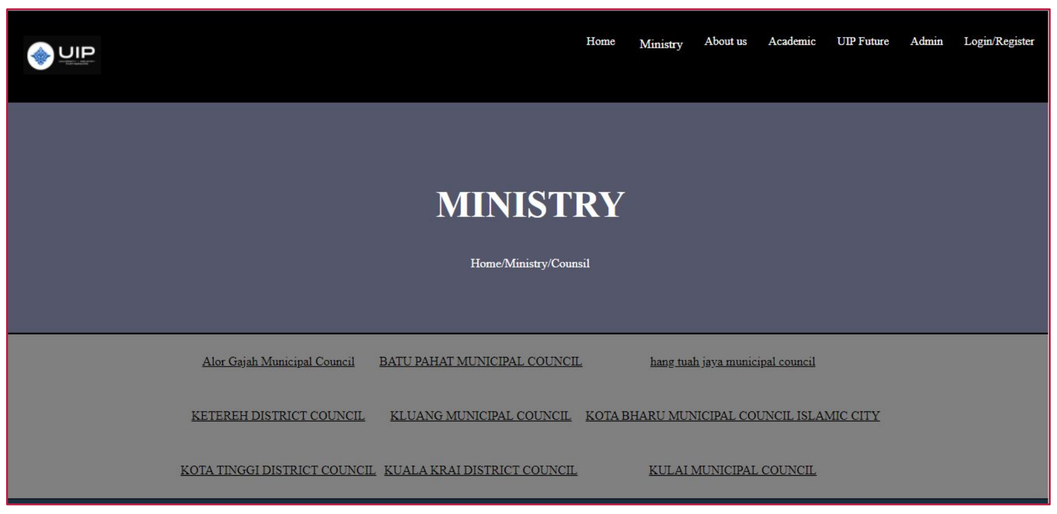

Figure 2: Ministry Module Webpage

Contribution

This project is the first ever database that centralize whole academia and government information into one. It can contribute for further researches and government projects and reduce the time finding the best person in it.

Conclusion

In the nutshell, initial objective in this project are achieved and base of the design in the project were built. However, more advance feature or design need to use to refine the project. Fortunately, every output required for this project are successfully shown out to the users.