Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

PROJECT TITLE: WEB SCRAPING AND DATA ANALYTICS (RANKING, SORTING AND STUDENT ACTIVITIES) FOR UIP DATABASES

ABSTRACT

In this globalized era, graduating from University is like a standard procedure for getting a diligent job in society. University also getting more in numbers and became common for students to achieve their high-level studies in every industry. However, information from every university's website are rarely standardized and some even fragmented which had caused the third party such as students and parents had their hard time to do researches with every single university. These phenomena will bring students into an unfavourable situation where they might choose the wrong university for themselves due to the difficulties in investigating the aspects and benefits of each university.

This project is aimed at developing an online platform for educators in universities to convert their areas of expertise into prototypes, curriculum, industry-friendly collaboration models and develop new areas of research with AI-enabled engines. This project will discover each educator and his/her specialization and achievement, then create a personalized AI expert system.

PROBLEM

With these loads of universities in Malaysia, it takes time for third parties such as teenagers, parents, and lecturers to explore and do research of the universities with every aspect such as the courses, subject material, social activities, portfolios, and et cetera.

However, some of the website information is fragmented and outdated such as the alumni of the university and facilities developments. Hence, potential students are unable to decide on their best choice of university to pursue their dream courses.

Constructive suggestions and feedback from all parties including students are vital for improvement purposes.

Fragmented data pose difficulties in the collaboration among industries, universities, and investors. Therefore, this project will also be looking into this aspect.

Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

Some universities produced numerous research outputs and non-commercialized patents that need more engagement and adoption by the industries. Moreover, technology is getting more contact with the methods in teachings of educators which change the norms in students learning experience. Hence, industries need to have a strong connection with the educators to ensure the enhanced learning scope and contents to be learnt by the students in universities.

SOLUTION

UIP, an AI-as-a-Service (AIaaS) platform will be built for educators in universities to convert and commit their areas of expertise into prototypes, curriculum, industry-friendly collaboration models and develop new areas of research with an AI-enabled engine.

UIP will be created and designed to showcase every select UIP research with various categories into one single platform. Research projects will be categorized as Agriculture, Board Reporting, Business, and et cetera. In the platform, users will be able to explore the UIP projects based on their interests.

The proposed 4 phases of the project are as follows:

1.1.1 Phase 1 Research and Prototypes

Various research is selected and prepare to upload into the platform to have a chance to collaborate with industries and develop a research prototype from the research itself. The researchers will have the opportunity to co-own the prototype when it is developed.

1.1.2 Phase 2 Website Scrapping Module

During this phase, we will scrap all information we need from the universities, government agency and industries.

1.1.3 Phase 3 Data Segmentation Module

In this phase, we will be gathering all data we got from phase 2 while we create a centralized online database to incorporate all the data into it. The raw data we collected from phase 2 will also convert into various sets of meaningful data clusters.

1.1.4 Phase 4 Dashboard Module

At this phase, we will create dashboards to export the data that can assist the users. Before being viewed by the specific users, the data will be analyzed by various methods to display out meaningful insights in different formats such as bar graphs, top-ranking charts, and et cetera.

Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

TARGET MARKET

The target market of this project is the universities researchers, the academic community, centres of excellence of universities, government agencies and ministries, SMEs, and corporations in 15 industry groups, investors, and undergraduates, and postgraduate students.

1. University researchers and the academic community

   They play the role of contributing and sharing their research or patent ideas to the platform as the platform tends to accept as much as possible of research and patents to display and commercialize the research output from education to the public.

2. Centres of Excellence of Universities

   This unit of the university could assist to filter out the best team of experts to help out with the research or the uncommercialized patents.

3. Government agencies and Ministries

   The agencies and ministries can plan their projects for future collaboration with industries and higher institutions for a better wisdom level in the country.

4. SMEs and Corporations in 15 industry groups

   Small Medium Enterprise plays a role of invest in various patents to be part of the contributors and use the patent technology or ideas in the community for the increment of their business and economics.

5. Investors (venture capitalists and private equity groups)

   We need investors as the platform is a long-lasting project and platform for the education community. Cannot be denied that a minority of the research or patients need long-lasting equity to continue and succeed in their research.

6. Undergraduates and postgraduate students

   The platform enables the students to find their interesting research topics from the platform and build up their passion from it.

Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

COMPETITION/CONTRIBUTIONS

The UIPFuture platform is mainly contributing to academic researchers and lecturers in every university that is collaborating with.

MILESTONES

The milestones for this project are as follows:

Timeline	Milestone		Milestone Goals
16/03/2021 - 30/03/2021	Web Scraping	Universities Staff Module	Scrap all the related information and produce output in .csv format.
25/05/2021 - 29/05/2021		Government Agencies Module	
19/04/2021 - 30/04/2021	UI/UX development	UIPFuture website module	Create a website as a platform to display various research.
		UIP Project module	Gather and categorized research projects.
	Database development	Central Database development	Push all the scraped information into one central database with auto search and filter functions.
		UI/UX for database	Development of UI/UX for the database to produce information needed and data flowchart.

Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

	Dashboard Design and Development	RoboAdvisor dashboard module	Design and development of the dashboard of the RoboAdvisor AI-based system.

REFERENCES

1. MyFinB Group, 2021, RoboAdvisor Module Specification, unpublished photograph, viewed 17 March 2021

2. MyFinB Group, 2021, Technical Stack Across System Components, unpublished photograph, viewed 17 March 2021

3. Bishop, R., 2021, Why Alumni Are Important for Higher Ed Institutions | Signal Vine. [online], Accessed 17 March 2021, https://www.signalvine.com/alumni/why-alumni-are-important-for-higher-ed-institutions#:~:text=Alumni%20serve%20many%20valuable%20roles,and%20career%20opportunities%20to%20students.

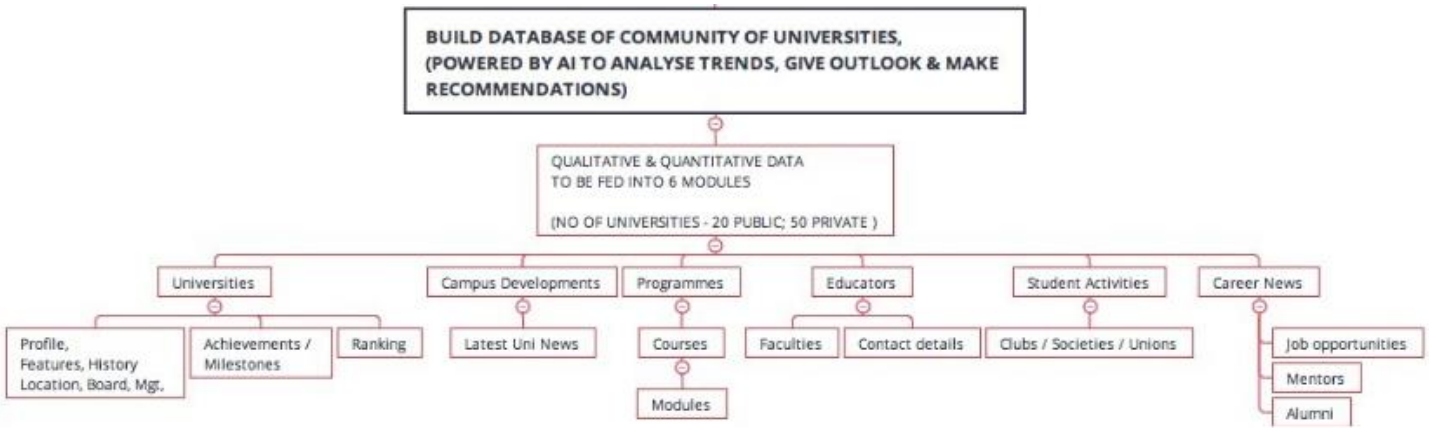
4. Love, J., 2021, [online] Leaderonomics.com. , Accessed 17 March 2021, https://www.leaderonomics.com/articles/personal/necessity-competition-leadership

5. Augustyn, A., 2021, supply and demand | Definition, Example, & Graph. [online] Encyclopedia Britannica, Accessed 17 March 2021 https://www.britannica.com/topic/supply-and-demand

6. Statista, 2021, Malaysia: higher education enrollment | Statista. [online], Accessed 17 March 2021, https://www.statista.com/statistics/794845/students-in-public-higher-education-institutions-by-gender-malaysia/#:~:text=Students%20in%20public%20higher%20education,Malaysia%202012%2D2019%2C%20by%20gender&text=In%202019%2C%20around%20291.53%20thousand,enrolled%20in%20public%20higher%20institutions.

Kong Mun Jun
Bachelor of Computer Science in Data Science
+6012-2874119
kongmj-wm18@student.tarc.edu.my
Dr Sandy Lim Siew Mooi

APPENDIXES



Appendix 1: Areas of information to collect in university

Technical Stack Across System Component			
Database	Scraping	Platform & Analytic	Dashboards/ UIUX
Oracle	Python	Python	Python
MySQL		Kotlin	Visual Studio
		Java	Google Colab

Appendix 2: Tools to use during development