Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	09 October 2023
Team ID	NM2023TMID09416
Project Name	Data Dominators : A Comparative study of Top Global Universities in Data Analytics

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

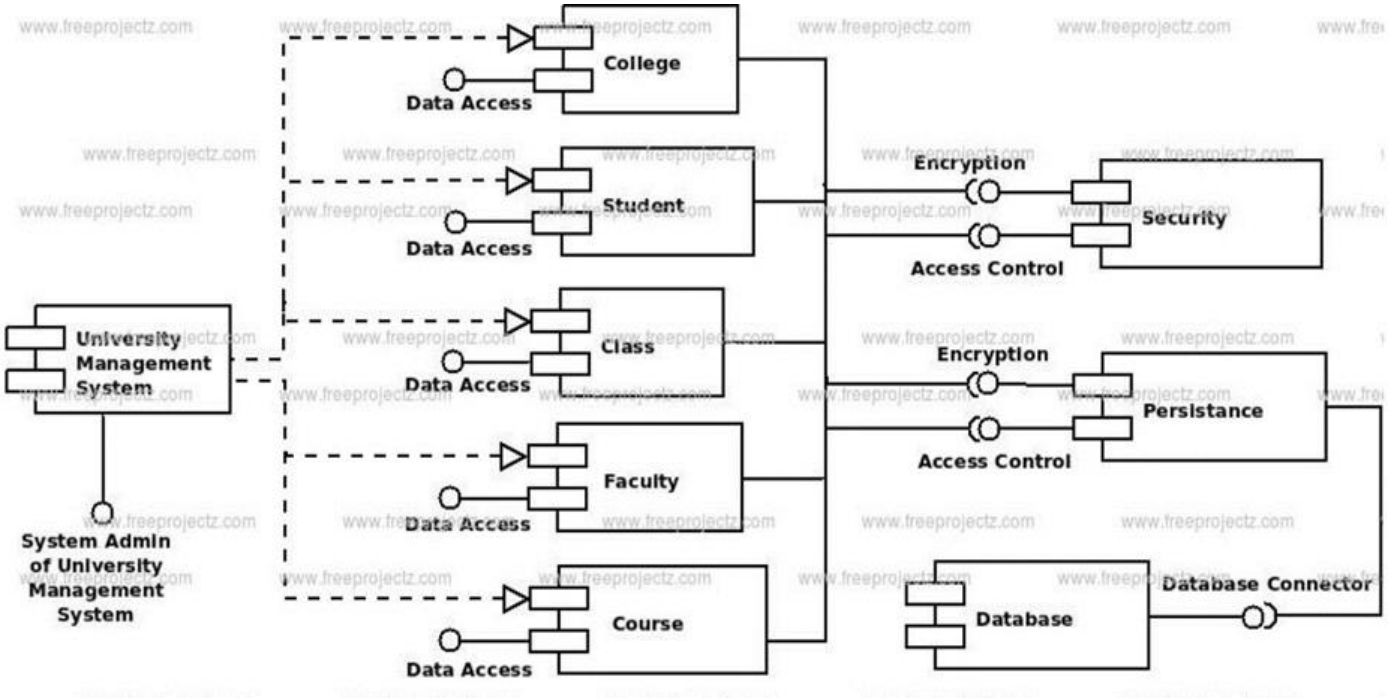

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	Web-based front-end for user interaction and data display	HTML, CSS, JavaScript /React etc.
2.	Back-End Application	Server-side logic for data retrieval and processing	Node.js, Express, Python, Django
3.	Authentication & Authorization	Secure user authentication and role-based access control	JWT, OAuth, Identity Providers
4.	Data Collection	Methods to gather data from various sources	Web Scraping, APIs
5.	Data Processing & Analysis	Analyze data, perform statistical analysis, and modeling	Python, R, Machine Learning
6.	Data Storage (Study Data)	Store structured data for the comparative study	Relational Database (e.g., SQL)
7.	Data Storage (User Data)	Secure storage of user information and feedback	NoSQL Database (e.g., MongoDB)
8.	Data Warehouse	Centralized data storage and retrieval for analysis	Amazon Redshift, Google BigQuery
9.	Data Visualization & Reporting	Create interactive visualizations and reports	D3.js, Chart.js, Plotly
10.	Content Management System (CMS)	Manage and update study content	Custom CMS, WordPress, Drupal
11.	Hosting & Deployment	Deploy web application and databases to cloud servers	AWS, Google Cloud, Heroku

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	User-Friendly Interface	The application should provide an intuitive and easy-to-use interface for users to interact with the study, select universities, and view the results.	HTML, CSS, JavaScript, React
2.	Data Accuracy and Reliability	Ensuring that the collected data is accurate and reliable is crucial for the credibility of the study. Data sources and integration processes should be carefully validated	Data Cleaning Tools, Python, R
3.	Scalability and Performance	The application should be able to handle a growing user base and large datasets efficiently. It needs to be scalable to accommodate future data analytics and user demands	Cloud Hosting (IBM, Google Cloud), Data Warehousing (Redshift, BigQuery)

S.No	Characteristics	Description	Technology
4.	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.)	Technology used
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Technology used