Geo Mapping and Analysis of Alumni Data (with Full Stack Web Development)

Redalia D'Souza
Department of Electronics and Telecommunication,
Don Bosco Institute of Technology
Mumbai, India.
redalia1998@gmail.com

Roni Thomas
Department of Electronics and Telecommunication,
Don Bosco Institute of Technology
Mumbai, India.
ronithomas10@gmail.com

Amanda Joseph
Department of Electronics and Telecommunication,
Don Bosco Institute of Technology
Mumbai, India.
mandyjoseph08@gmail.com

Aishwarya Kadam
Department of Electronics and Telecommunication,
Don Bosco Institute of Technology
Mumbai, India.
ishwaryaanil@gmail.com

Namita Agarwal
Department of Electronics and Telecommunication,
Don Bosco Institute of Technology
Mumbai, India.
namita.dbit@dbclmumbai.org

Abstract—**A dedicated alumni network is considered to be a great asset to any institution. Alumni are the face of the institution, those who represent the institution in the real world. With more people looking to upskill, many alumni may prefer their alma mater for higher courses or certificate programs. Institutions can also keep track of courses popular among their alumni. Alumni provide the best form of marketing for their alma mater, i.e, word of mouth marketing, making alumni relationships a critical component of developing the institution's brand. A large number of institutions turn to donations from their alumni to subsidize their undergraduate programs [13]. The main aim of this work is to build an interaction between alumni, administrators and the students; a system that will be able to collect and analyse alumni dataset of an institute, to provide trends, locate anonymously and provide easy access to the same.**

Keywords—alumni, database, geo mapping, analysis

I. BACKGROUND

Alumni website is created for the students that have graduated from the institution. This is an online website that allows former students to take advantage of the benefits and services that an institution offers after graduation. The alumni network is becoming important in the development of the institution because of their vast potential that benefits both the institution and the students. According to [8], alumni relations are important to provide feedback and counsel, generate financial support, participate in governance and assist with student recruitment, both directly and indirectly. Mapping of alumni worldwide not only defines the level of prestige of the institute, but also facilitates in identification of favoured career paths chosen by those now accomplished in their respective fields. The core of Geographic Information Systems (GIS) are digital maps. By connecting databases containing geo locations, various multi-layered maps can be created with themes such as traffic, weather, availability of resources, etc. Relational databases are noted to have additional functionality compared to NoSQL databases, when it comes to geospatial applications, as mentioned in [6].

II. MOTIVATION AND OBJECTIVE

A. Motivation of the Project

1) Alumni of a college generally stay in touch with their immediate friends but find it difficult to stay connected with other college mates.

2) Creating an engaged, supportive alumni network is crucial to an institution's success. If communication stops once graduates leave an institution, their understanding of the university will become stale. Instead, they should be kept informed so they can remain engaged and keep abreast on the progress of the university.

3) Contact between alumni can be used to forge business connections and to gain references or insight in a new field.

4) An alumni network can benefit the institution in many ways such as mentoring students, organizing alumni days, having training sessions during alumni days, arranging work practices, and proposing topics of theses, to raise funds for the organization.

5) It is important to carry out good follow-up marketing of alumni events. Also mapping them on sites gives an added benefit for people to look up to the institution.

B. Objective of the project

1) To analyse alumni data with respect to employability, higher studies, etc.

2) Extract valuable information to use in strategic making, trend analysis, etc.

3) To engage alumni and current students in all aspects of the university.

4) To provide a user-friendly website that visualises the analysed data.

III. LITERATURE SURVEY

In website development and data analytics, there are many paths and tools to achieve the same goal. For instance, there are a plethora of database management services provided by various organisations. In order to make an informed decision over an optimum course of action, papers on the following topics were surveyed:

- Database
- Website Design
- Cloud Services
- Geo Mapping

It has been inferred from paper [1] that MariaDB is being used with a standard query language called SQL whereas MongoDB has its query language named Mongo Query Language. MariaDB offers much more additional

functionality to the MySQL engine. In paper [2], the author states that RDBMS databases are scalable, secure, easy to configure, and access and support ACID properties for transactions. RDBMS also supports OLAP systems which provide the capability for complex calculations, trend analysis, and sophisticated data modeling, thus providing added functionality in the geospatial domain. A MySQL database might not meet the requirements of availability, scalability, quick data backup and data recovery, while NoSQL databases are inconsistent and inefficient for handling complex queries [8]. Paper [3] mentions that Facebook API uses RESTful protocol and responses are in JSON format. LinkedIn uses Javascript API and REST API, the latter of which uses OAuth 1.0a to authorize users and begin making REST API calls using any programming language. The technologies used in this work are included in the Services Web Mapping era, as mentioned in paper [4]. In [5], which is a review of [4], it is stated that API interfaces unlock online data servers and provide easier access to store, retrieve and link multiple online data repositories. The API used in this project is the Mapbox API.

IV. METHODOLOGY

The core components of implementation revolve around research, data collection, database creation, web designing and analysis and mapping.

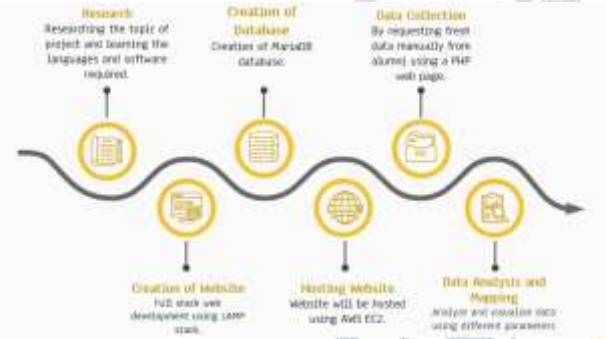

Fig 4.1: Methodology

1. *Research*

Collection of all the insights required to go about with the project which includes research papers, learning the languages, exploring the software requirements, locating the materials & making notes.

2. *Creation of Website*

The website would act as a room to showcase all the required data and visualization. It will also give the user an in-house user experience bringing everything under one roof. A website greatly enhances the data visualization experience, enhancing dull mathematical representations into visually pleasing infographics [12].

3. *Creation of Database*

A database is a collection of related data stored off a computer and organized to enable information to be retrieved as needed. Many cloud based applications provide Database As A Service (DBaas), which allow developers to seamlessly integrate their services. However, in cases such as AWS EC2, open source databases such as MariaDB are supported. Data is defined by the three Vs of Big Data, viz., volume, veracity and variety. in [7], it is mentioned that velocity should be considered when dealing with geo spatial data, as it is better to analyse the incoming, realtime data to perform dynamic analysis on it.

4. *Hosting Website*

A web hosting service is a type of Internet hosting service that allows individuals and organizations to make their website accessible via the World Wide Web. The preferred cloud service in this case is Amazon Elastic Cloud Compute.

5. *Data Collection*

Data collection is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes. One of the biggest problems faced by data analysts is the availability of an exhaustive dataset that contains sufficient information to provide statistically accurate outputs.

6. *Data Analysis & Mapping*

Data analysis is a process of inspecting, cleansing, transforming, and modeling data to discover useful information, informing conclusions, and supporting decision-making, also plotting graphs for easy visualization. The simplest form of internet mapping includes static maps, which simply include images (such as jpegs). Web based maps or maps built using APIs are far more interactive and allow developers to tweak its components in a way that would provide the best suited results[11].

V. IMPLEMENTATION AND RESULTS

The information has been provided by the college along with all the new information acquired after circulating the form, which is then analyzed and displayed on the website. There are two parts to Web Development the first being frontend which includes the design of the website or the designing interface of the web application using languages such as HTML and CSS. The backend includes server-side programming, i.e., it communicates the client interface with the database and the logic control using languages like PHP, JavaScript, Python, etc.

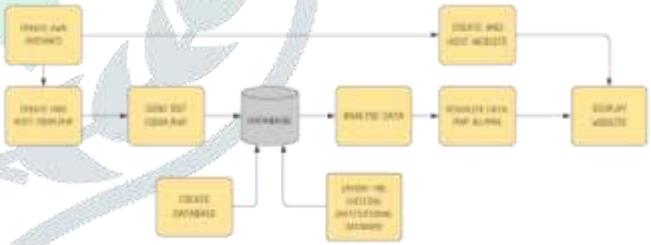

Fig 5.1: Block Diagram

1. *Create AWS instance*

AWS is a secure cloud service platform offering compute power, database storage, content delivery, and other functionality also running web and application servers in the cloud to host dynamic websites. An AWS EC2 t2.micro instance is used which is operated on Ubuntu (Linux) and is powered by 2 CPUs and has 8 GB of storage. The website is connected to the instance with the help of free, remote computing software, Mobaxterm.

2. *Create and host website and form.php*

Website is created using the LAMP (Linux Apache MySQL PHP) stack. The form is written in PHP and is connected to a MariaDB “alumni” database A free web-based vector graphics editor and prototyping tool, Figma, is used for UI development. The website is then hosted on the above-mentioned EC2 instance to be accessed by users.

3. Creation of Database

The database used in this project is MariaDB, a community-developed, open-source, commercially supported fork of the MySQL relational database management system. Data will be loaded from the form into the database.

4. Data Analysis, Visualisation, and Mapping

The stored data will be analyzed and visualized using Metabase. This analyzed data will also be displayed on a dedicated webpage on the website. Visualization can also be done using Python libraries such as pandas and NumPy. Alumni will be mapped by fetching coordinates of the cities they are currently based in and plotting them on a map API such as Mapbox.

5. Display Website

The final website containing widgets, APIs, and dedicated web pages will be hosted on the EC2 instance. The website will also support user authentication using OAuth, an authentication protocol that allows users to approve one application interacting with another on their behalf without giving away their password.

Results:

The data visualization in this section is an analysis of approximately 50 entries. As the dataset grows the accuracy of the analysis will increase. Fig(6.5) shows the percentage of alumni in different cities. It is noted that a majority of the alumni currently reside either in or around Mumbai or in different parts of India. This could also mean that most alumni from the current department prefer to take up jobs than go abroad for their higher studies.

Fig(6.4) shows the graph of how CGPA affects the choice of specialization in higher studies. It is noted that among those students who prefer higher studies, students having higher CGPA in their undergraduate course feel more confident to go for a degree in STEM such as Masters in Engineering and Masters in Science. Those with a slightly lower CGPA preferred to do a Masters in Business Administration and those with a considerably lower GPA went for a Masters in Management.

Fig 6.1: Form

Fig 6.2: MySQL Database

Fig 6.3: Mapping

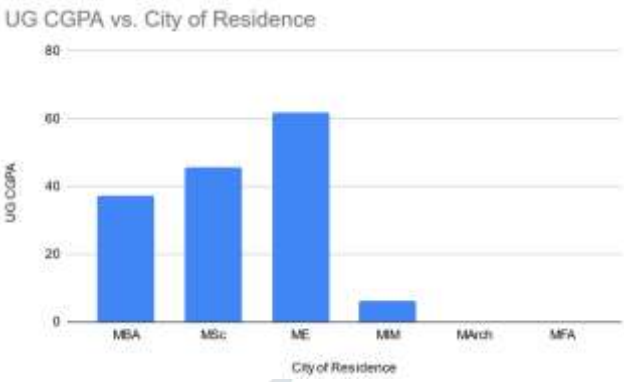

Fig 6.4: Graph of UG CGPA vs PG Specializati

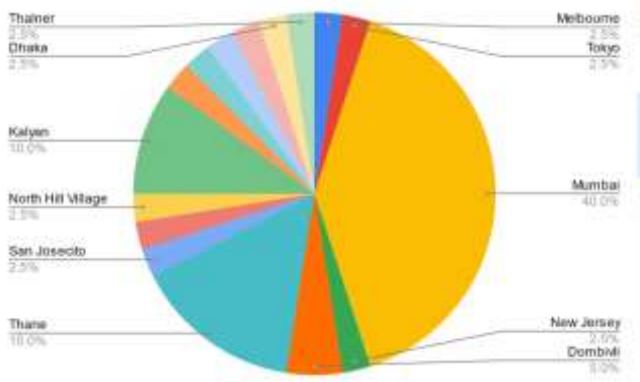

Fig 6.5: Percentage of alumni in different cities

CONCLUSION

The outcome of this work would help the students to be up to date with the current practical world not only in terms of selecting the appropriate field for further studies, but will also help them acquire information about the current jobs, internships, industries, and their requirements. This would be achieved first by circulating a form to collect the required information from the alumni, to build a dataset, and to perform a detailed analysis on the same. After analyzing the data, the above-analyzed information would be displayed in the form of graphs, pie charts, and tables which will help in the identification of trends concerning higher studies, placements, etc. which will, in turn, provide insights to students currently studying to acquire the required information and make informed decisions. Results and output of mapping will be displayed and stored on the website.

REFERENCES

[1] S. Lakshmi Devi, P. S. Vijayalakshmi, "A Comparative Study: MariaDB Vs MongoDB", Volume 7 Issue 1, January 2018, International Journal of Science and Research (IJSR), Paper ID: ART20179511, DOI: 10.21275/ ART20179511.

[2] Waleed Al Shehri, "CLOUD DATABASE AS A SERVICE", International Journal of Database Management Systems (IJDMS) Vol.5, No.2, April 2013, DOI: 10.5121/ijdms.2013.5201.

[3] Gurudatt Kulkarni, Ramesh Sutar, Jayant Gambhir, "Cloud Computing-Infrastructure as a Service: Amazon EC2", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 1, Jan-Feb 2012, pp.117-125.

[4] Bert Veenendaal, "ERAS OF WEB MAPPING DEVELOPMENTS: PAST, PRESENT AND FUTURE", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B4, 2016.

[5] Bert Veenendaal, Maria Antonia Brovelli, Songnian Li, "Review of Web Mapping: Eras, Trends and Directions", ISPRS Int. J. Geo-Inf.2017,6, 317; doi:10.3390/ijgi6100317.

[6] Elena Baralis, Andrea Dalla Valle, Paolo Garza, Claudio Rossi, Francesco Scullino, "SQL versus NoSQL Databases for Geospatial Applications", IEEE International Conference on Big Data (Big Data), 2017, DOI:10.1109/BigData.2017.8258324.

[7] Jae-Gil Lee, Minseo Kang, "Geospatial Big Data: Challenges and Opportunities", Big Data Research, vol. 2, no. 2, pp. 74 – 81,2015, visions on Big Data.

[8] Xi Zheng, "Database as a Service - Current Issues and Its Future", 2018 ACM. ISBN 978-1-4503-2138-9.DOI:10.1145/1235.

[9] Tarun Kumar, Yeeshu Prateek, Prajwal Atharga, Dr. Rajashekarappa, Prof. V. K. Parvati, 2019, Alumni Database Management System, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCRACES – 2019 (Volume 7, Issue 10).

[10] "Alumni Relations" [Online] Available: https://jcu.edu/about-us/administrative-offices/university-advancement/departments/alumni-relations

[11] "Web Mapping: An Introduction" [Online] Available: https://mapasyst.extension.org/web-mapping-an-introduction/

[12] "10 Websites with Data Visualization Driving User Experience" [Online] Available: https://designmodo.com/data-visualization/

[13] "Why Alumni Are Important for Higher Ed Institutions" [Online] Available: https://www.signalvine.com/alumni/why-alumni-are-important-for-higher-ed-institutions