Data Visualization for Multimedia Super Corridor Malaysia using Dot Distribution Maps

Mohd Suffian Sulaiman[1], Amira Yasmin Raimey[1] and Azri Azmi[2]

[1]Universiti Teknologi MARA, Malaysia {suffian@tmsk.uitm.edu.my, amira.yasmin0510@gmail.com}
[2]Universiti Teknologi Malaysia, Malaysia {azriazmi@utm.my}

ABSTRACT

Multimedia Super Corridor (MSC) is a Malaysian government designated zone developed as a catalyst for the advancement of the information and communication technology (ICT). Currently, the locations of each MSC Malaysia status companies are represented in a list. Thus, it is time-consuming and inconvenient for users to go through a list of 3213 of MSC Malaysia companies. This study proposed to develop a webpage application to visualize the locations of MSC Malaysia companies in Klang Valley using dot distribution maps technique. Google Maps Javascript API is used to display the map in the webpage application. The data of the companies are collected and reorganized with additional information. The webpage application also utilizes markers as symbols to display the exact location of the MSC Malaysia companies. Besides, the tooltips feature is also included. Pan and zoom control features are provided by Google Maps Javascript API. From the result of system testing, it shows that the webpage application functions, work properly and the locations of MSC Malaysia companies are presented in a more understandable way in which ease for users to search for the exact location. By applying dot distribution maps technique, the users can easily view and identify the locations of MSC Malaysia companies in Klang Valley.

Keywords: Data Visualization, JavaScript, Multimedia Super Corridor, Dot Distribution Maps, Google Maps, API

I   INTRODUCTION

A structured and unstructured of large and complex data are referred as big data in which the traditional processing techniques are unable to operate on these data (Husin & Abdullah, 2019; Taylor-Sakyi, 2016). The characteristics of big data are 5V's which consist of volume, velocity, variety, veracity and value (Hilda, Srimathi, & Bonthu, 2016). According to International Data Corporation (IDC), the collection of data is growing at an exponential rate (Olshannikova, Ometov, Koucheryavy, & Olsson, 2015). A massive volume of data is complex to be analyzed, therefore the data visualization techniques will be used to solve this issue. Data visualization is vital in visualizing the

big data efficiently (Razak, Fauzi, Gining, & Jamaluddin, 2019; Sulaiman, Mohamad, Hanafi, & Wahab, 2018). The visualized data using the data visualization techniques are more understandable as they are presented in graphical or pictorial form (Sadiku, Shadare, Musa, & Akujuobi, 2014).

In this study, MSC status companies are chosen as the dataset domain. MSC is a Malaysian government designated zone developed as an accelerator for the progression of the information and communication technology (ICT). MSC Malaysia companies are initiative projects by the fourth Malaysian Prime Minister, Tun Dr. Mahathir Mohamad. MSC Malaysia companies are classified into four different clusters of ICT focus area which are Creative Content Technologies (CCT), Global Business Services (GBS), Institutes of Higher Learning (IHLs) and Information Technology (InfoTech). The advancement of ICT is important to the society to transform Malaysia into a knowledge-rich economy by the year 2020 to achieve the Vision 2020 (Injau, 2011). MSC Malaysia is committed to nurture and guide the talent groups from primary and secondary students, undergraduates, fresh graduates and working professionals, digital maker, and premier IHL. This study is proposed to focus on fresh graduates from local universities. It is a job opportunity for the fresh graduates because MSC Malaysia is a great platform for the fresh graduates to pursue their career in digital technology to enhance their knowledge in ICT area.

Based on our study, we found that the user faces difficulties to identify the location and search the MSC Malaysia company on the current web site. It is due that the user can only search for the name of the MSC Malaysia company. Alternatively, the user can search for each company by scrolling up and down through the list. However, it is very time-consuming for the user to identify the location of the company because the user can only view the address once clicked on the company's name to view more details information of the company. Currently, the MSC Malaysia company has 2,954 active companies. The data stored for each company includes the company name, core activities, address and contact and these data are still presented as a text list. This presentation of data is inefficient to be analyzed.

Data visualization can be categorizing to five types which are temporal, multidimensional, hierarchical,

network and geospatial(Ward, Grinstein, & Keim, 2015)(Wilke, 2019). In this study, the geospatial or spatial data visualizations type is chosen because it is suitable category in order to display the real-life physical locations of the MSC Malaysia companies in Klang valley.

In geospatial category, there are several techniques that can be used to visualize the data such as cartogram, choropleth map, proportional symbol map, dot distribution map and dasymetric map. Cartogram is map that represents a mixture of statistical and geographical data in the geographical sections are adjusted in which the statistics of the population and income are presented (Nusrat, Alam, & Kobourov, 2018). Cartogram can be categorized into four different types which are contiguous cartograms, non-contiguous cartograms, dorling cartograms and rectangular cartograms. Choropleth map is used to visualize the weather condition or vulnerability data of region. Choropleth also is the most common technique used to depict the quantitative thematic geodata. The spatial units such as countries, states and cities are frequently used in choropleth map to emphasize the differences of the geographically distributed data by color or patterns (Skowronnek, 2015). Proportional symbol map is a map that visualize the location data such as countries, states or regions which varies by the variable attributes and the locations data are presented using circle, square, bar or a pictographic symbol. The advantage of using proportional symbol map technique is the attribute values of the measured location data and conceptual location data are represented together (Kunigami, Rezende, & Souza, 2014). Dot distribution map also known as density map uses dots in visualizing the location data. It is a technique used in representing the point-based data in which each point displayed on the map represents data with a known geo-location or a value aggregation (Polisciuc, Alves, & Machado, 2015). Dasymetric map is used in assigning population data to display the spatial data and its auxiliary data. The outcomes of dasymetric mapping can be validated by evaluating the relationship between the density of population and the auxiliary data using the statistical linear regression (Liu, Peng, Wu, Jiao, & Yu, 2018).

Different techniques of geospatial can be applied on the particular scenario and case study. Therefore, the dot distribution map is chosen since it is suitable with the objective of this study to visualize the MSC Malaysia companies in Klang valley.

II RESEARCH METHOD

This section will focus on the process of developing the data visualization for MSC Malaysia status companies. The waterfall model (Sulaiman & Ghani, 2011) was chosen as the software process model to represent the comprehensive study process as depicts

in Figure 1. This model is used because it works well for smaller projects where the requirements are well understood.
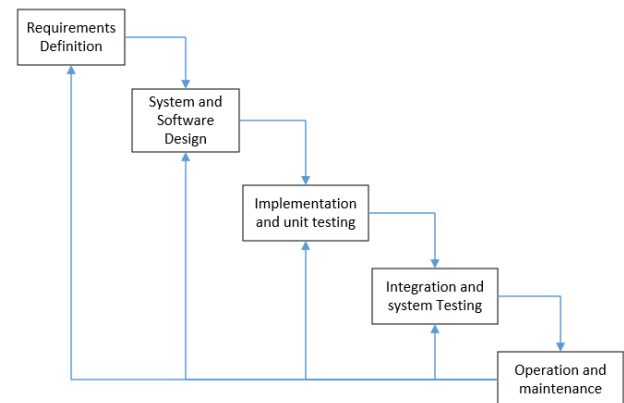

Figure 1. Waterfall model

In requirements definition phase, the system's services, constraints and goal is established by consultation with target user and literature review from selected website, books and articles from journal and conference proceeding. The data of active MSC Malaysia companies were collected from Malaysia's Open Data Portal (Jamil, n.d.). The data extracted from Malaysia's Open Data Portal is in .xlsx format as shown in Table 1. The extracted data consists of 3213 active MSC Malaysia companies. It consists of 370 CCT cluster, 605 GBS cluster, 2119 InfoTech cluster and 119 IHLs and Incubators cluster.

Table 1. Sample Of Extracted Data From Malaysia's Open Data Portal

File ID	Company Name	Year Of Approval	Main Cluster	CoreActivities
2	C3/3/2789	UGT GLOBAL (MALAYSIA) SDN. BHD. (F	2006	Global Business Services	1. Development of software solutions and applications. 2. Maintenance and testing of softw
3	C3/3/9536	APIQ Consulting Sdn Bhd	2016	Global Business Services	Establish and operate a Global Business Services Center rendering 1) IT Outsourcing in:i. Syste
4	C3/3/6275	ZAM Sdn Bhd	2008	InfoTech	Research, development, and commercialisation of IT applications & solutions to the R
5	C3/3/689	Gii Vision Sdn Bhd	2001	InfoTech	Development of e-Farmasi manager, application software for the retail pharmacies as well
6	C3/3/10492	Biz Excellence Systems Sdn Bhd	2017	InfoTech	Research, development and commercialization of the following solution: 1. BizExcel e-Gove

Figure 2 shows the use case diagram (Sulaiman, 2003) representing the functional requirement of data visualization for MSC Malaysia companies (Sulaiman, Rosli, & Ghazali, 2009). The target user can be an internship student, fresh graduates or any qualified local or international people.
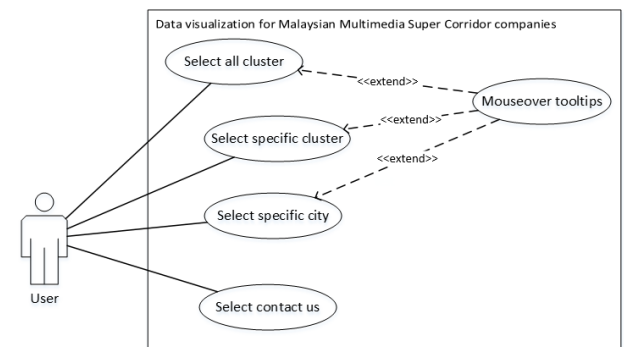

Figure 2. Use Case Diagram

In system and software design phase, an overall system architecture model is proposed as depicted in Figure 3. The architecture partitions the requirements to either hardware or software systems (Gani & Sulaiman, 2005; Sulaiman, Tamizi, Shamsudin, & Azmi, 2020). All the data are stored inside the web server. The user need to use any browser in order to access the data visualization for MSC Malaysia companies. In implementation and unit testing phase, it is realized by using related software and hardware such as notepad++ as programming editor for Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript, Google Chrome as web browser, XAMPP as a web server and laptop with Windows 10 operating system.
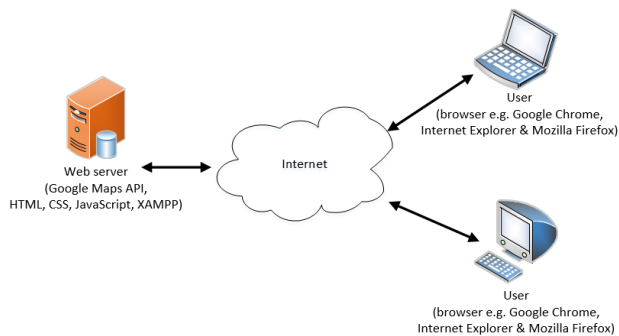

Figure 3. System Architecture

In this study, the Google map is used because Google is a top map service provider in the world. The Google map is added by loading the application programing interface (API) of the map from a specified URL as shown in Figure 4. It shows that the URL's parameter contains the API key to execute the initMap() function.

```
<script async defer
      src="https://maps.googleapis.com/
maps/api/js?key=AIzaSyAYtoJbLNNoaoofa5Q
Hq2mam-WPBwmlm70&callback=initMap">
</script>
```

Figure 4. Initiate Map Code Excerpt

Figure 5 shows the initMap() function. In this function, the coordinates of Selangor state that includes with latitude and longitude were obtained from the Map Coordinates apps. This coordinate will be the center of the map.

```
function initMap(){
var map = new
google.maps.Map(document.getElementById
('map'),{zoom: 10.2,
      center:new
google.maps.LatLng(3.0738, 101.5183)
   });
}
```

Figure 5. Initmap() Function Code Excerpt

The array list is utilized to store the data companies which include the company's name, cluster, contact, website, address, latitude, longitude and main activities. All the data will be sorted according to its cluster and city. Figure 6 shows a sample code excerpt using initCyberjaya() function.

```
function initCyberjaya(){
   var locations8 = [['Xhai Studios Sdn
Bhd', 2.922712, 101.661302, "Creative
Content Technologies", "+6018-2862168
(T)", "info@xhaistudios.com",
"xhaistudios.com", "i-Tech Tower, Jalan
Impact, Cyber 6, 63000 Cyberjaya,
Selangor", "Provision of pre-production,
production and post production services
for: 1) Xhai Mobile Games Development,
Provision of implementation, maintenance
and technical services related to the
above services 2) Research, development
and commercialization of the following
solution: Gamer's Diary(GD), Provision
of implementation, maintenance and
technical services related to the above
solution"],];
}
```

Figure 6. Initcyberjaya() Function Code Excerpt

The marker is used as a symbol to indicate the exact location of MSC Malaysia companies. The coordinates for the marker to be placed will be retrieved from an array list named locations. Once the coordinates have been retrieved, the marker for each company will be generated. Figure 7 shows the excerpt code of the marker.

```
for(var i=0; i<locations.length; i++){
      var loc = locations[i];
      var marker = new
google.maps.Marker({
          position: {lat: loc[1],
lng: loc[2]},
          map: map
}
```

Figure 7. Marker Coding Excerpt

The discrete piece of information of the company which includes company name, address, contact, core activities and URL website is generated using tooltips. The tooltips will display when the user hover their mouse to the particular marker. Figure 8 shows the tooltips coding excerpt. As shown in Figure 8, the name of company is stored at loc[0], address stored at loc[5], contact that consists of telephone number is stored at loc[3] while email is stored at loc[4] in the array list. Core activities is stored at loc[6] and website is stored at loc[7] in the array list.

```
title: "Company Name" +"\n" + loc[0] +
    "\n\nAddress\n" + loc[5] +
  "\n\nContact\n" + loc[3] + "\n"  +
  loc[4] + "\n\nCore Activities\n" +
  loc[6] + "\n\nWebsite\n" + loc[7]
```

Figure 8. Tooltips Coding Excerpt

The drop-down menu features on the dot distribution map enable the users to display different datasets on the map. Figure 9 shows the drop-down menu for cluster which consists of CCL, GBS, InfoTech and IHLs and Incubators. This drop-down menu enables user to display the MSC Malaysia companies according to selected cluster. For example, if a user clicked on CCT, the application will only display the MSC Malaysia companies that belong to that cluster. This similar approach will be applied for cities in Klang Valley such as Petaling Jaya, Shah Alam, Cyberjaya, Subang Jaya, Bangsar and Kuala Lumpur. All the functions, variables and reserved words are tested in unit tesing phase to confirm that the code is doing things right (Azmi & Ibrahim, 2011a).

```
<div class="dropdown">
    <button class="dropbtn"> Clusters
</button>
        <div class="dropdown-menu">
            <a onclick='initMap()'>All
Clusters</a>
            <a
onclick='initCCTMap()'>Creative Content
Technologies</a>
            <a
onclick='initGBSMap()'>Global Business
Services</a>
            <a
onclick='initITMap()'>Information
Technology</a>
            <a
onclick='initIIMap()'>IHLs &
Incubator</a>
        </div>
</div>
```

Figure 9. Drop-Down Menu Coding for Cluster Excerpt

In integration and system testing phase, the functional testing is done to ensure all the functional requirements as stated in the requirements definition phase were verified (Azmi & Ibrahim, 2011b). The last phase, operation and maintenance is ignored since it is beyond of our study scope due to the main objective is to develop the data visualization using dot distribution map (Sommerville, 2015).

III RESULTS AND DISCUSSION

The data visualization for selected 80 MSC Malaysia companies are presented in the graphical user interface as illustrated in Figure 10. By a single click on map button, the web page will display all the selected companies from all clusters. It will ease the target user

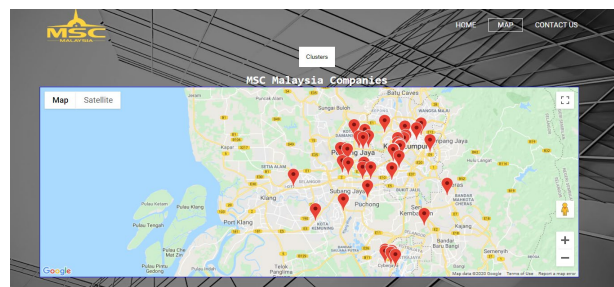to identify the location and distribution of the companies.

Figure 10. Marker for Selected MSC Malaysia Companies

A. Map based on cluster

Figure 11 illustrates the location of MSC Malaysia companies based on CCT cluster. The user can select an appropriate cluster by using the drop-down menu. The user can select another cluster such as GBS, InfoTech and IHLs & Incubator as well.
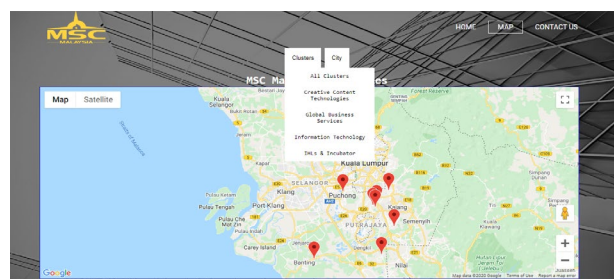

Figure 11. Drop-Down Menu For Selected Cluster

B. Map based on city

Figure 12 shows the drop-down menu for selected companies in particular city in Klang Valley such as Petaling Jaya, Shah Alam, Cyberjaya, Subang Jaya, Bangsar and Kuala Lumpur.
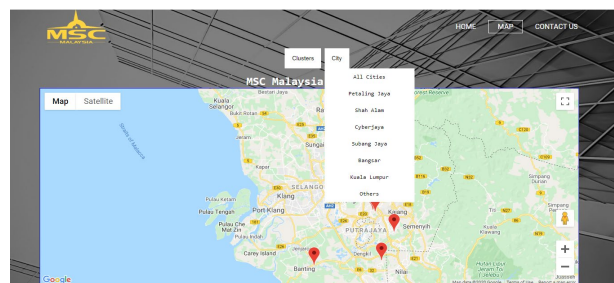

Figure 12. Drop-Down Menu for Selected City

C. Tooltips

Figure 13 shows the tooltips that will display the information such as company name, cluster name, contact, address, core activities and website URL when the user mouseover to the particular marker.

The list of MSC Malaysia companies previously were shown in MSC Malaysia directory on the Malaysia Digital Economy Corporation Sdn Bhd (MDeC)'s official website (Malaysia Digital Economy Corporation, n.d.). However, the current official website no longer displays the list of active MSC companies. It only shows the number of active companies which are 2954 as of May 2019. The data collected from the Malaysia's Open Data Portal listed an active MSC companies as of year 2017 consists of 3212 companies. Thus, it is difficult to determine whether the MSC companies in 2017 are still active in 2019. In addition, the current data do not provide the complete company profile such as contact number, email, address and coordinate of the company's location. Therefore, the data need to be updated when the new current active MSC Malaysia companies are obtained.
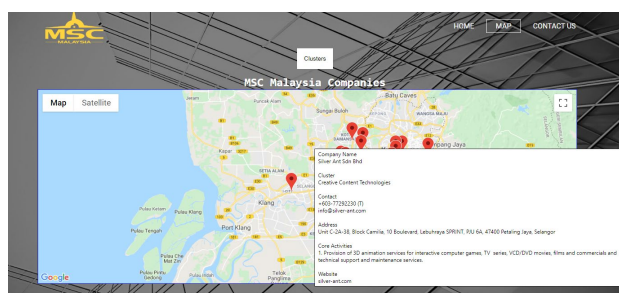

Figure 13. Tooltips for Each Companies

IV CONCLUSION

This paper presents the development of data visualization for MSC Malaysia status companies using dot distribution maps technique. The target users for this study is for fresh graduates that are looking for job or any qualified local or international people that would like to pursue a career whether for full time, part time or an internship with MSC Malaysia companies located especially in Klang Valley. This study covers all clusters of the MSC companies which are Creative Content Technologies (CCT), Global Business Services (GBS), Information Technology (InfoTech) and Institutes of Higher Learning (IHLs) and Incubators. This study can be further improved by adding searching feature that will enable the users to search the nearest location of MSC Malaysia companies based on city or state.

ACKNOWLEDGMENT

REFERENCES

Azmi, A., & Ibrahim, S. (2011a). Formulating a Software Traceability Model for Integrated Test Documentation: a case study. *International Journal of Information and Electronics Engineering*, *1*(2), 178–184. https://doi.org/10.1115/1.859735.paper78

Azmi, A., & Ibrahim, S. (2011b). Implementing test management traceability model to support test documents. *International Journal of Digital Information and Wireless Communications*, *1*(1), 109–125. https://doi.org/10.1017/CBO9781107415324.004

Gani, M. K. A., & Sulaiman, M. S. (2005). *Object Oriented Component Based Framework For Web Transactional Application.*

Hilda, J. J., Srimathi, C., & Bonthu, B. (2016). A Review on the Development of Big Data Analytics and Efective Data Visualization Techniques in the Context of Massive and Multidimensional Data. *Indian Journal of Science and Technology*, *9*(27). https://doi.org/10.17485/ijst/2016/v9i27/88692

Husin, N. D. I., & Abdullah, N. A. S. (2019). Overlapping issues and solutions in data visualization techniques. *Indonesian Journal of Electrical Engineering and Computer Science*, *16*(3), 1600–1608. https://doi.org/10.11591/ijeecs.v16.i3.pp1600-1608

Injau, H. (2011). *Evaluation of Multimedia Super Corridor (MSC Malaysia) Contribution in Malaysian Economy*. Ritsumeikan Asia Pacific University. https://doi.org/10.16194/j.cnki.31-1059/g4.2011.07.016

Jamil, H. M. (n.d.). data.gov.my. Retrieved February 16, 2020, from http://www.data.gov.my/data/dataset/active-msc-malaysia-status-companies

Kunigami, G., Rezende, P. J. de, & Souza, C. C. de. (2014). Optimizing the layout of proportional symbol maps: Polyhedra and computation. *INFORMS Journal on Computing*, *26*(2), 199–207. https://doi.org/10.1287/ijoc.2013.0557

Liu, L., Peng, Z., Wu, H., Jiao, H., & Yu, Y. (2018). Exploring urban spatial feature with dasymetric mapping based on mobile phone data and LUR-2SFCAe method. *Sustainability (Switzerland)*, *10*(7), 1–15. https://doi.org/10.3390/su10072432

Malaysia Digital Economy Corporation. (n.d.). MSC Malaysia Directory. Retrieved February 27, 2020, from https://mdec.my/what-we-offer/msc-malaysia/

Nusrat, S., Alam, M. J., & Kobourov, S. (2018). Evaluating Cartogram Effectiveness. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1105–1118. https://doi.org/10.1109/TVCG.2016.2642109

Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data*, *2*(1), 1–27. https://doi.org/10.1186/s40537-015-0031-2

Polisciuc, E., Alves, A., & Machado, P. (2015). Understanding Urban Land Use through the Visualization of Points of Interest. In *Proceedings of the Workshop on Vision and Language* (pp. 51–59). https://doi.org/10.18653/v1/w15-2810

Razak, M. Z. A., Fauzi, S. S. M., Gining, R. A. J., & Jamaluddin, M. N. F. (2019). Data visualisation of vehicle crash using interactive map and data dashboard. *Indonesian Journal of Electrical Engineering and Computer Science*, *14*(3), 1405–1411. https://doi.org/10.11591/ijeecs.v13.i3.pp1405-1411

Sadiku, M. N. O., Shadare, A. E., Musa, S. M., & Akujuobi, C. M. (2014). Data visualization. *Educational Measurement: Issues and Practice*, *33*(2), 3–4. https://doi.org/10.1111/emip.12034

Skowronnek, A. (2015). Beyond choropleth maps: A review of techniques to visualize quantitative areal geodata.

Sommerville, I. (2015). *Software Engineering* (10th Ed.). Pearson.

Sulaiman, M. S. (2003). *Development of Job Applying System for KUTKM Using UML*. Universiti Teknologi Malaysia.

Sulaiman, M. S., & Ghani, M. K. A. (2011). *Software Requirement, Analysis & Design, Object Oriented Approach with UML*. Penerbit Universiti, Universiti Teknikal Malaysia Melaka.

Sulaiman, M. S., Mohamad, A. T., Hanafi, H. F., & Wahab, M. H. A. (2018). Data Vizualization for Computer Inventory. In *Knowledge Management International Conference*.

Sulaiman, M. S., Rosli, M. M., & Ghazali, J. N. (2009). Development of Mobile Learning Application for Enhancement of Teaching Tools. In *The 2nd International Conference on e-Learning (iCEL)*.

Sulaiman, M. S., Tamizi, A. A., Shamsudin, M. R., & Azmi, A. (2020). Course recommendation system using fuzzy logic approach. *Indonesian Journal of Electrical Engineering and Computer Science*, *17*(1), 365–371.

Taylor-Sakyi, K. (2016). *Big Data: Understanding big data. CIM Magazine* (Vol. 11). https://doi.org/10.4018/978-1-5225-9750-6.ch001

Ward, M., Grinstein, G., & Keim, D. (2015). *Interactive Data Visualization*. AK Peters/CRC Press.

Wilke, C. O. (2019). *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures. O'Reilly Media.* O'Reilly Media.