# Big Data Visualization:
## Application in Visualizing Learning Activities

Thanh Tam Nguyen

School of Business/IT
James Cook University Australia
Singapore Campus, Singapore

thanhtam.nguyen@my.jcu.edu.au

Insu Song

School of Business/IT
James Cook University Australia
Singapore Campus, Singapore

insu.song@jcu.edu.au

**Abstract.** Information visualization helps users quickly identify interesting and significant events and patterns from data that are otherwise too detailed or complex to discern. For instance, rich content and convenience provided by social networks allow students to exchange ideas and collaborate. In this context, students are interested in exploring learning activities of other students without having to read through lots of textual contents. Students tend to have the interest of finding information concerning their majors, contents of the subjects and their co-learners. However, with the huge amount of information available, it is difficult to identify relevant information for the users. In this paper, we highlight challenges and opportunities in the visualization of big data and learning activities on social network and present a new method of visualizing learning activities of students on social networks. The method allows students to view trends in content and activities near them and around the world. A new application is developed based on the method and is evaluated. The results show that students find it fun and easy to explore learning activities of their peers.

**Keywords:** Big Data Visualization, Data Visualization, Social Networks, Education, eLearning, Learning Activities

## 1 Introduction

Visualization helps in boosting human's cognitive processes [1-4]. It has been proved to offer a range of advantages to the process in which people acquire knowledge, which include (1) understanding of large data, (2) unexpectedly interesting ways of perceiving information, (3) quick recognition of errors and outliers in

adfa, p. 1, 2011.

data set, (4) identification of patterns in data, and (5) ease at hypotheses formation out of the data [5].

With the mentioned benefits, information visualization has been widely deployed in many areas, from scientific study, crime analysis to business and education. Hence, a wide range of techniques and methods of enhancing visualization representation and interactivity have been studied and developed over the last decades. Some noteworthy studies include: visualization of unstructured temporal data with parallel rendering algorithm [6], taxonomies of interaction techniques [7], the focus+ context technique [8], treemaps for visualizing hierarchical data structure while making use of all of the available space [9], and artificial reality in visualization [10].

With the remarkable growth in data generation, visualization plays a more essential part than ever in facilitating the process in which people obtain insights into data [11]. The total amount of data generated is expected to experience a significant growth, surpassing approximately 40 zeta bytes in 2020, and it will be increasing by 40% each year in the next decade [12]. Nonetheless, the study showed that only a small portion of the collected data was tagged (around 3%) and the data that were analyzed was even less (about 0.5% of the world's digital data). This necessitates approaches to representing data in a more intuitive way so that data is ready for users to easily perceive. Visualization is expected to help in solving the challenge.

The same situation occurs in education. Knowledge has had much shorter life than before [13]. The application of data visualization in education, visualizing learning activities particularly can help improve learning analytics significantly. However, the process to choose what information is to be shown and the manner of displaying the information so that learners can find it fun and relevant have remained a challenge.

In this paper, we propose an approach of visualizing learning activities in social networks using a 3D scene platform. First we review existing literature to identify gaps and opportunities in improving the learning analytics through data visualization. We then develop a new model of visualizing learning activities using 3D data visualization framework. We evaluate its effectiveness in making the learning process more fun and motivating and effective. Questionnaires and interviews are used for the evaluation. Results show that our method of visualization of learning activities in social network made learning more fun and easier. Additionally, it also indicates that the model helps improve the students' engagement and motivation on the subjects.

## 2 Background

### 2.1 History of Data Visualization

Data Visualization came into the world since pre-17th century in the form of early maps. It then evolved through several periods, from new graphic forms, modern graphics to current state of high dimension interactive visualization [14-16]. The milestone project for data visualization history came up with one significant finding, which is the root of graphical portrayal was the entwinement of several factors includ-

ing statistical graphics, thematic cartography, data visualization, statistical thinking and developments in technology.

### 2.2 Definition of Data Visualization

Different studies have come up with various descriptions of data visualization [17-21]; and there is no single universally uniform definition of this concept. As stated by [17], data visualization could be considered as a process in which data was taken as input and visual representations were generated as output. The output had to be displayed in a way so that users could easily understand and quickly grasp the insight into the data. Gershon et al. [20] defined information visualization as an operation where data, information and knowledge were transformed into visual forms so that the innate visual ability of people's could be made use of. Ware [21] gave a summary of the benefits given by visualization, which include: faster understanding of large datasets, unpredictable perspectives of seeing information, easier recognition of outliers and errors, efficient identification of patterns within data, and easier hypothesis formation process.

### 2.3 Big Data

According to [22] and [23], in spite of being used commonly all over the world, the term "big data" remained to be an abstract concept which different people had dissimilar views on its definition. However, basically, big data could be considered as any data sets that were so big or complex that it was impossible for traditional hardware or software applications to acquire, perceive, manage and process within an acceptable time.

Big data has been often described as having 3 main features. As stated in his article, [24] used 3Vs (Volume, Variety, and Velocity) to talk about the characteristics of big data. Volume meant that the amount of data being generated was growing exponentially. Variety indicated the various types of data created, which include structured data, semi-structured data and unstructured data as well. Velocity meant that the creation of data and data analysis had to be carried out in a fast manner.

### 2.4 Big Data Visualization Application Areas

Big data visualization has been deployed in various fields, and a great number of benefits have been offered due to its use. Those areas cover from scientific study, biology, crime analysis to business and education. Table 1 gives a summary of some interesting applications of Big Data Visualization.

A survey was conducted at the IEEE Information Visualization 97 Symposium by Wright [25] to give a review on the application of Information visualization in business. The findings from the survey showed that visualization added value to businesses in many ways. For example, exploratory analysis facilitated the understanding and conceptualization of complicated issues. Visual representation could lead to faster

user cognition which improved the decision making process. In addition, communication gaps could be shortened thanks to data visualization.

In education area, information visualization could offer several advantages when being applied for educational purposes. Teachers and course conductors could use the generated graphical data to analyze the learning activities of the learners to achieve an overview of the students' learning. In addition, using data visualization in learning process could encourage and motivate learners more so that their studies might be more productive [26]. Naps et al. [27] ran the Working Group on Improving the Educational Impact of Algorithm Visualization, carrying out online discussions and online survey as well as personal meeting in order to determine how visualization could contribute to making improvements in education and computer science. The research provided results of educational benefits that visualization offered; for example, 90% of the respondents asserted that the teaching experience is more enjoyable; 86% agreed that visualization improved the level of participation of students, and 83% of them thought with visualization, class became more fun for their students.

### 2.5 Big Data Visualization in Learning Network

As mentioned above, the world's digital data has been being massively generated, and the same situation is occurring in education area. The Web has made it easier for people to create, organize and share information, experience and knowledge. That brings out the need for setting up learning networks where learners can connect and improve productivity of their learning process. However, as suggested in [28] the abundantly available information also resulted in overloaded information. Learners may find it hard to spot the important information that they want. Visualization is expected to help learners in grasping the overview of their learning environment and easing the learning process.

Visualization does play a significant role in learning analytics as well. Many researches have been conducted to examine learning analytics. For instance, different sources of observations were integrated and analyzed to give a better understanding of learning activities and discover new learning contexts [29,30]. A research was working on evaluating the usage of learning systems and how visualization facilitate the process of obtaining insights into the learning activities [31 ]. Experiments of a collaborative e-learning system were built up to discuss how to analyze tracks which were created by communication tools and how it could be helpful for the re-engineering purposes [32].

## 3 Methodology

By reviewing the previous and existing literature on online learning platform, visualization for educational area, we identify gaps in literature and figure out opportunities. According to our literature review, not much research has been done in visualizing the learning activities within the social network.

We develop a model of visualizing learning activities within a learning network to examine if this approach would help in increasing the learning effectiveness and encouraging collaborative learning. Within the scope of this research, we create a survey for learners who use social network for their studies in order to identify their demands, requirements and problems. An online questionnaire is created with the content consisting of these indicators: age, major, frequency of social network usage for education purpose. With the purpose of supporting students' learning activities, the questionnaire examines user experience to determine the usability of 3D learning activities visualization in terms of engagement, motivation, encouragement and collaborative. Target users of the survey are JCU students who are using KOPO MES, an online learning platform, for their studies. After that, the collected questionnaire data are analyzed to rationalize the utility and usability of social learning activities visualization model.

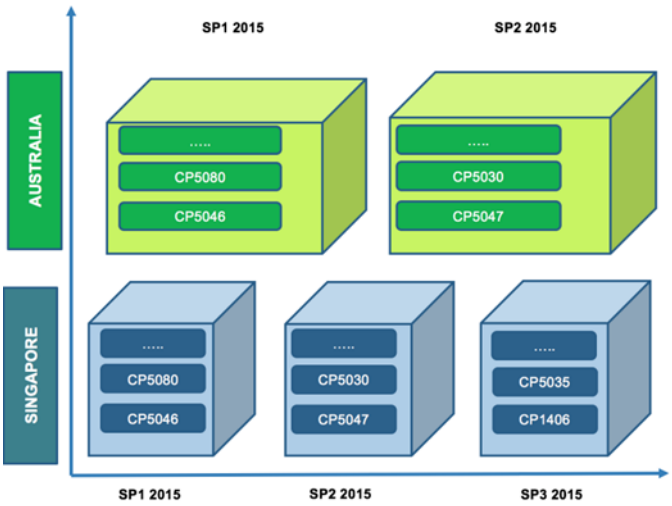


Figure. 1. Demonstration of the learning activities involved in KopoMES. There are two dimensions: location and time. The horizontal axis represents time (semesters). The vertical one represents locations. Each box represents a learning entity including students studying in the same location and time.

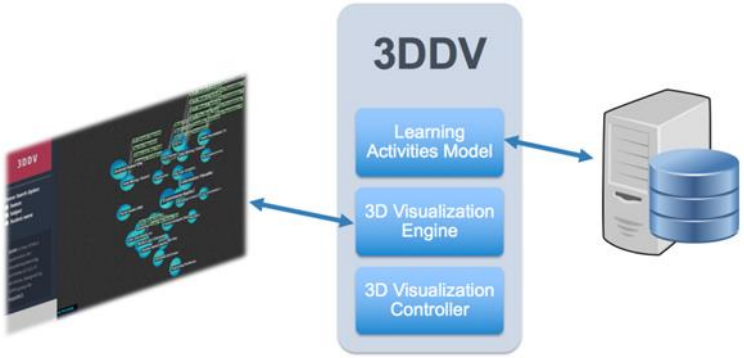


Figure 2. The 3DDV (3D Data Visualization) system is an HTML5 web application used to visualize the learning activities on any mobile devices. The system comprises of the learning activity model which is used to retrieve learning activities data from a remote server through the provided web service (a.kopo.com), the 3D visualization engine (view), and 3D visualization controller.

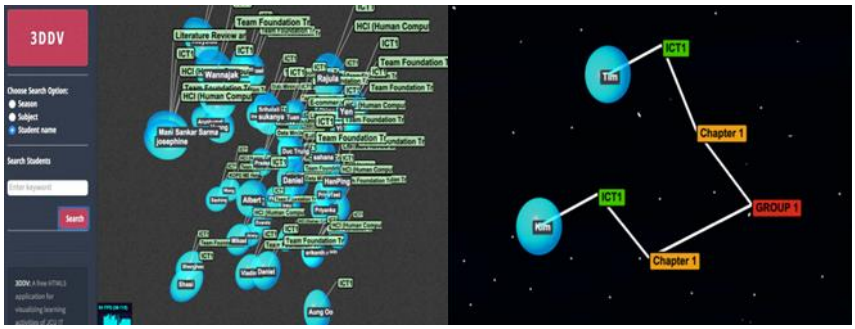


Figure 3. Visualization of students' learning activities. A student is represented by a sphere and subjects and groups are presented by labels. Standard touch interface on mobile devices and search tools can be used to explore learning activities of other students in specific time and location. Figure 3(b) shows that two students Tim and Kim enrolling in the same group for subject ICT1

Figure 1 demonstrates the learning space of JCU students (in both Singapore and Australia). Students can take part in one or more learning activities in various locations and time. For example, Figure 1 illustrates two locations: JCU Australia and JCU Singapore campus over one year period (3 tri-semesters). Students can choose to be transferred between campuses and same subjects offered in the two campuses but at different timing. This is because JCU Australia runs two-semester system while JCU Singapore runs a tri-semester one.

Students are often found to be interested in other students' learning activities such as the subjects they are studying, location and time. Our system aims to visualize these activities so that students can help each other forming tight student networks.

Figure 2 explains the proposed system for visualizing the learning activities. Users use the 3D Visualization Engine as an interface to interact with the system; the controller will send the request to the server through a web service to retrieve the requested data and filled into the learning activities model component. An open-source 3D library, namely ThreeIS, is used to generate all 3D objects and scene on users' browser.

Based on conducting literature review and survey, we develop a model of visualizing the learning activities in social network using HTML5, CSS3 and JavaScript. The JavaScript 3D library called ThreeJS is utilized. Figure 3 illustrates that each student is presented as a sphere and connected to her learning activities. Students who have the same learning activities are connected together.

## 4 Experiment Setup and Results

In order to evaluate the usability of the new system, we created a set of synthetic data comprising of 48 users and 27 subjects. This data is used to generate the test database for the visualization system of 3DDV.

33 participants were presented with demo and given the access to the 3D visualization system to use the system and have an overview about the topic. The seventeen of them were tested with an aim to find out the new model gives them motivation and a better overview of the subjects. Sixteen participants were asked to interact with the system, exploring the subjects they are to see whether the system improves their engagement, and helps them solve their problems. Finally, a questionnaire is given to the participants to fill in. Indicators included in the questionnaire consist of: easiness, use-friendliness, motivation, collaboration. Also, they were asked to give comment on improving the system.

In order to prove to usability and utility of 3DDV, we evaluated the approach using the questionnaire. The USE Questionnaire (Usefulness, Satisfaction, and Ease of Use Based) is used for this survey. The content of the survey included 2 parts: Basic information of participants and the usefulness of 3DDV with 5 Likert Scale (Strongly agree, Agree, Neutral, Disagree, and strongly disagree). The survey's target users are JCU students, more focusing on KOPO MES users.

Over 83% of the participants agree that they have problem in finding relevant information about courses, subjects and solutions online. This shows a strong demand for 3DDV and the prospective of the system. There is one fact that many of the students have not known the way of using social network effectively for their studies.

There are 60.87% of participants from 18 and 25 years old, 30.43% of them are from 26 to 31 years old. The majority of participants are at 21-30 age groups. Males take more than 60% out of 33. The variety of participants' nationalities helps increase the accuracy of the results. They all agreed with the hypothesis that had been raised in the research.

## 5 Conclusion

In this paper, by reviewing previous literature, we identified application of big data visualization, especially in learning activities. Our research shows that learners have an interest in using social network when performing their studies with the aim to figure out who are taking which subject, where and when.

A new model of visualizing learning activities is proposed, namely learning space. We then developed a system for realizing the model using 3D visualization framework. The results show that students find it intuitive to explore other students' activities, otherwise rather complex and time consuming to navigate manually. We believe that in the future visualization will play a significant part in education area. Learning will be much easier and fun when an easy-to-use learning activity explorer and social media are integrated.

## References


1. Bauer MI, Johnson-Laird PN (1993) How diagrams can improve reasoning. Psychological science 4 (6):372-378
2. Larkin JH, Simon HA (1987) Why a diagram is (sometimes) worth ten thousand words. Cognitive science 11 (1):65-100
3. Mayer RE, Gallini JK (1990) When is an illustration worth ten thousand words? Journal of educational psychology 82 (4):715
4. Card SK, Mackinlay JD, Shneiderman B (1999) Readings in information visualization: using vision to think. Morgan Kaufmann,
5. Colin W (2004) Information visualization: perception for design. San Francisco, CA: Morgan Kaufmann
6. Ma K-L, Stompel A, Bielak J, Ghattas O, Kim EJ Visualizing very large-scale earthquake simulations. In: Supercomputing, 2003 ACM/IEEE Conference, 2003. IEEE, pp 48-48
7. Yi JS, ah Kang Y, Stasko JT, Jacko JA (2007) Toward a deeper understanding of the role of interaction in information visualization. Visualization and Computer Graphics, IEEE Transactions on 13 (6):1224-1231
8. Lamping J, Rao R, Pirolli P A focus+ context technique based on hyperbolic geometry for visualizing large hierarchies. In: Proceedings of the SIGCHI conference on Human factors in computing systems, 1995. ACM Press/Addison-Wesley Publishing Co., pp 401-408
9. Johnson B, Shneiderman B Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In: Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on, 1991. IEEE, pp 284-291
10. Erickson T (1986) Artificial realities as data visualization environments: Problems and prospects. Virtual Reality-Applications and explorations:3-22
11. Heidemann J, Willinger W (2014) Internet visualization. Computing 96 (1):1-2

12. Gantz J, Reinsel D (2012) The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC iView: IDC Analyze the Future 2007:1-16
13. Siemens G (2014) Connectivism: A learning theory for the digital age.
14. Friendly M, Denis DJ (2001) Milestones in the history of thematic cartography, statistical graphics, and data visualization. U RL http://www datavis ca/milestones
15. Friendly M (2005) Milestones in the history of data visualization: A case study in statistical historiography. In: Classification—the Ubiquitous Challenge. Springer, pp 34-52
16. Chen C-h, Härdle WK, Unwin A (2007) Handbook of data visualization. Springer Science & Business Media,
17. Wills G (2012) Visualization toolkit software. Wiley Interdisciplinary Reviews: Computational Statistics 4 (5):474-481
18. Ellis G, Dix A (2007) A taxonomy of clutter reduction for information visualisation. Visualization and Computer Graphics, IEEE Transactions on 13 (6):1216-1223
19. Fayyad UM, Wierse A, Grinstein GG (2002) Information visualization in data mining and knowledge discovery. Morgan Kaufmann,
20. Gershon N, Eick SG, Card S (1998) Information visualization. interactions 5 (2):9-15
21. Ware C (2012) Information visualization: perception for design. Elsevier,
22. Ward JS, Barker A (2013) Undefined by data: a survey of big data definitions. arXiv preprint arXiv:13095821
23. Chen CP, Zhang C-Y (2014) Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences 275:314-347
24. Beyer M (2011) Gartner says solving big data challenge involves more than just managing volumes of data. Gartner. 2011) http://www gartner com/it/page jsp
25. Wright W (1998) Business visualization adds value. IEEE Computer Graphics and Applications (4):39
26. Romero C, Ventura S (2010) Educational data mining: a review of the state of the art. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 40 (6):601-618
27. Naps TL, Rößling G, Almstrum V, Dann W, Fleischer R, Hundhausen C, Korhonen A, Malmi L, McNally M, Rodger S Exploring the role of visualization and engagement in computer science education. In: ACM SIGCSE Bulletin, 2002. vol 2. ACM, pp 131-152
28. Hiltz SR, Wellman B (1997) Asynchronous learning networks as a virtual classroom. Communications of the ACM 40 (9):44-49
29. Avouris N, Fiotakis G, Kahrimanis G, Margaritis M, Komis V (2007) Beyond logging of fingertip actions: analysis of collaborative learning using multiple sources of data. Journal of Interactive Learning Research 18:231-250
30. Marty J-C, Heraud J-M, Carron T, France L (2007) Matching the performed activity on an educational platform with a recommended pedagogical scenario: A multi-source approach. Journal of Interactive Learning Research 18 (2):267-283

31. Mazza R, Milani C Gismo: a graphical interactive student monitoring tool for course management systems. In: International Conference on Technology Enhanced Learning, Milan, 2004. pp 1-8
32. Barre V, El-Kechaï H, Choquet C Re-engineering of collaborative e-learning systems: evaluation of system, collaboration and acquired knowledge qualities. In: 12th Artificial Intelligence in Education AIED, 2005. pp 9-16