Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Investigating Relevance of Prior Learning Data Connected through the Blockchain

Patrick OCHEJA[a*], Brendan FLANAGAN[b], Rwitajit MAJUMDAR[b] & Hiroaki OGATA[b]
[a]*Graduate School of Informatics, Kyoto University, Japan*
[b]*Academic Center for Media and Computing Studies, Kyoto University, Japan*
*ocheja.ileanwa.65s@st.kyoto-u.ac.jp

Abstract: Learners often change their learning environment over the course of their education. This makes it difficult to measure their engagement across different contexts due to a lack of seamless connection and shared analytics across heterogenous learning systems. Previous research has shown that access to prior engagement information of learners can be useful in enabling personalization, learning content design and early identification of problematic prerequisite topics. In this paper, we connect learning systems at two different schools through the blockchain to enable the transfer of learning footprints across both schools. Our primary aim is to investigate the relevance of students' prior engagement behaviour and provide stakeholders with actionable insights on dashboards. Specifically, we analyze the engagement behaviour in Junior High School grade 3 Math course of students who are currently in High School grade 1. Engagement in this context is defined based on five metrics: self-evaluation, cognitive behaviour, backtracking behaviour, time commitment and content completion rate. We further validate relevance by measuring the correlation between students' engagement level and their final score. Our analysis shows a significant difference in mean scores of very high and very low engagement students. Also, for each of the courses and scores, we provide stakeholders access to the learning materials used, assessments taken and the solutions by the students. Finally, we present implications for the field and present potential directions on how to use decentralized learner data to improve learning outcomes.

Keywords: Blockchain, education, engagement, learning behaviour, students, lifelong learning

1. Introduction

Teachers often face a common problem of not knowing the past learning engagements of their students. While final grades or scores may be contained in academic transcripts, it is difficult to measure students' engagement from transcripts. Trowler (2010) defines student engagement as the interaction between the time, effort and other relevant resources invested by students and institutes towards optimizing learning experience and to enhance students' performance. The differences in learning purposes, preferences, and motivations of students can result in different types of engagement behaviour during learning which may in-turn affect their performance (Li & Tsai, 2017). Previous research has shown that students' engagement in the learning environment is closely related to their learning outcome (Hu & Li, 2017; Lu, Huang, Huang & Yang, 2017). Thus, giving teachers access to their students past engagement could equip them with information about the possible challenges students may face, eliminate repetitive learning, how to adapt learning contents and provide support to students with prior low engagement.

To measure students' engagement at different times, it becomes necessary to access and analyze their total experience while learning at an institute. However, access to students' learning data after they change school is often difficult. This is largely attributed to the heterogeneous nature of learning systems and the lack of transferability of lifelong learning logs across schools (Baker, 2019). The advent of decentralized technologies such as the blockchain opens up new ways to address this problem. Ocheja, Flanagan, & Ogata (2018) proposed a blockchain of learning logs platform (BOLL) that can connect learning behaviour logs of students across different schools on a secure and immutable ledger. While the BOLL system solves the problem of learning data continuity, this paper presents a first of its kind research on providing teachers access to insights drawn from their students prior learning data such as engagement and learning outcome. In this work, we use the BOLL platform to provide teachers access to their students past learning engagements and investigate the relevance of students' past learning behaviour logs. For example, when students move from JHS 3 to HS 1, their HS 1 teacher is given access to the students' past learning behaviour logs. However, the teacher does not have data analytics skills to know if the learning behaviour logs have any effect on the final scores obtained. Our main argument is that it is not enough to provide access to past learning logs: the relevance of such data should also be communicated to the stakeholders. This is important because in most cases, stakeholders do not have the required data analytics to carry out such investigations on their own. We also provide a first of its kind access to the learning materials and assessment data (questions, students' and teachers' solutions) used by the student at their previous school using the marketplace (Boll-M) feature of Boll (Ocheja, Flanagan, & Ogata, 2019a). Specifically, this paper is focused on answering the following research questions:

RQ1. What are the engagement levels of students at a past learning environment?

RQ2. How relevant are these engagements to students' past learning outcome?

RQ3. How can teachers access additional information about learning outcomes?

The rest of this paper is organized as follows: In the second section, we review related works on student engagement analysis, investigate the correlation between engagement, and performance and highlight the originality of this work. The third section introduces our research methodology and the processes involved in retrieving the learning behaviour logs of students from two different schools on BOLL using the same blockchain identity, and how we calculate the engagement metrics. We present the results from our analysis and visualizations for stakeholders in the fourth section. Finally, in section five, we discuss the key findings of this research, open challenges, possible solutions and future work.

2. Related Work

Fredricks, Blumenfeld and Paris (2004) classified student engagement in three dimensions: Behavioural, Emotional, and Cognitive. Behavioural engagement entails students' participation in learning, academic tasks and school activities, positive conduct, and absence of disruptive behaviours (Fredricks et al, 2004). Emotional engagement deals with students' affective reactions in class such as interest, boredom, happiness, sadness, and anxiety (Connell & Wellborn, 1991; Skinner & Belmont, 1993). Cognitive engagement refers to students' motivation, effort and strategic use of provided learning resources through different methods such as self-regulation and meta-cognition (Fredricks et al, 2004). This work focuses on measuring students' behavioural and cognitive engagement from past and current learning behaviour logs of students across schools and platforms.

Most of the previous studies mainly investigated students' engagement using data from the current learning environment (Li & Tsai, 2017; Lu et al., 2017; Vytasek, Patzak & Winne, 2020). While these past studies have provided useful results for the problems they addressed, we argue that students' prior engagement can provide additional important information very early: solving the *cold-start problem*. This could help learning analytics systems to make more effective personalization decisions such as recommendations, and learning preference settings. Also, because students' prior engagement and achievement are predictive of their subsequent goals (Martin & Liem, 2010), providing teachers and students with such information becomes useful especially for teaching and self-regulation (Boekaerts & Corno, 2005). Access to learner data across different institutes they have attended is still lacking and this research makes the first practical effort to facilitate transfer of learner data across schools and measure its impact on teaching and learning outcomes.

Our unique contribution in this research is to make students' prior engagement accessible when students change school. Teachers at their new school can then access and use insights from such data to improve their teaching, students' engagement and learning outcome. It is important to note that this research does not propose a new method of measuring student engagement: our main focus is to use existing techniques to measure students' engagement based on their learning behaviour logs at their previous school.

3. Visualization of Prior Learning Data

We implemented four different visualizations for stakeholders to view the past engagement of students. The *learner profile* shown in Figure 1 gives a comprehensive summary of a student's past engagement and their achievements. This can also tell the teacher if the past engagement is correlated to the student's score or not. For each of the assessments, one can also view the student's solution as well as the correct solution. The *Engagement Transition* in Figure 2 gives stakeholders ability to view change in engagement level of a group of students using iSAT (Majumdar & Iyer, 2016). For example, teachers can check transition across a period of time to know when (or at what point in the past) a student's learning behaviour changed (improved or needs intervention). The teacher can also compare engagement changes across courses, contents or activities.

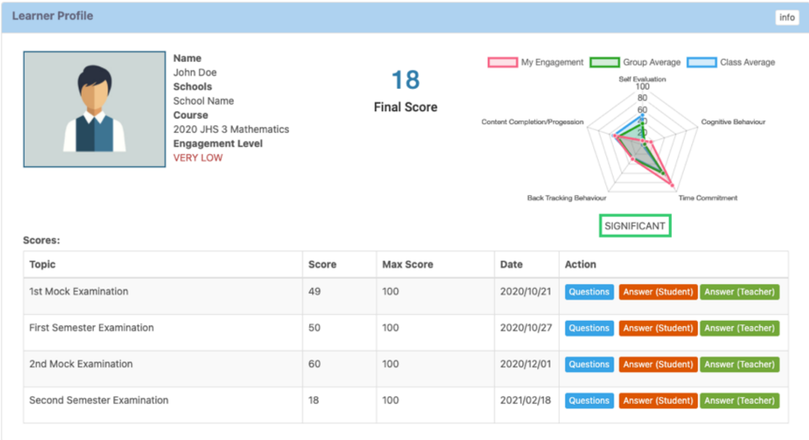


Figure 1. Learner Profile.

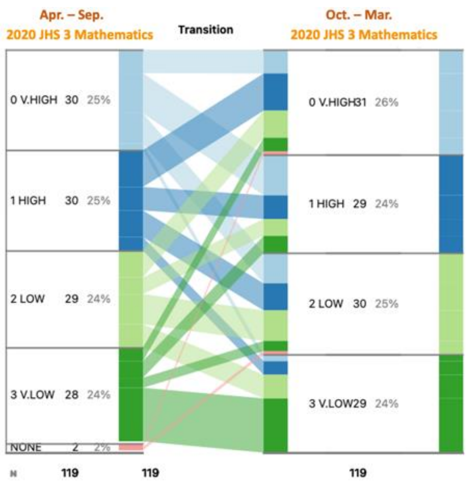


Figure 2. Temporal Change in Engagement Level.

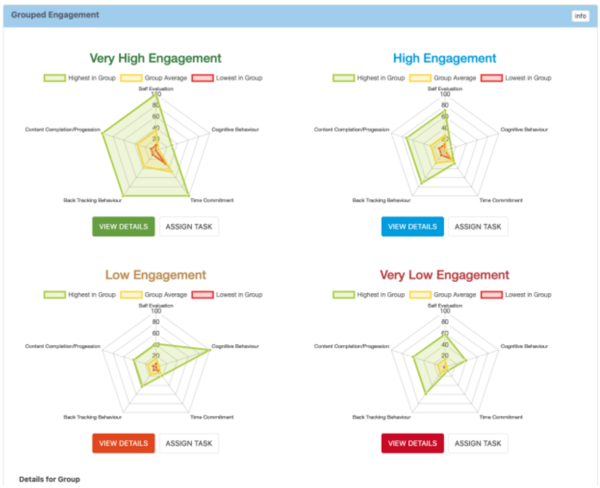
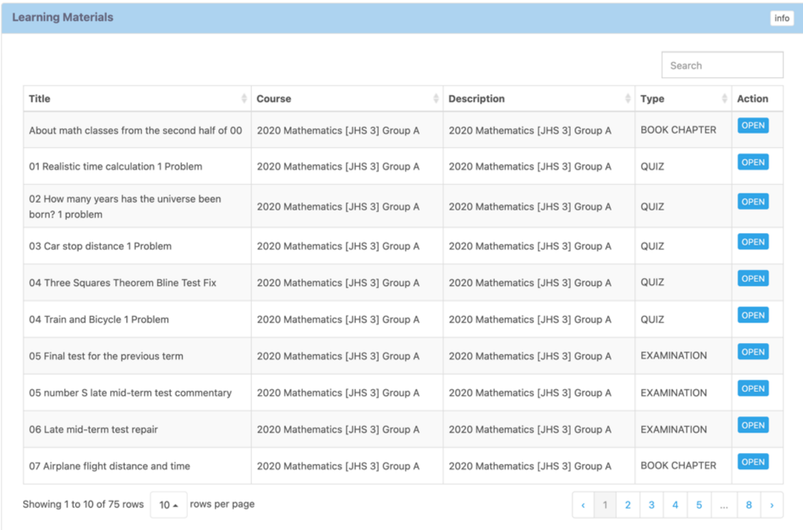


Figure 3. Detail profiles of Engagement Groups.

The *engagement groups* visualization in Figure 3 enables stakeholders to view engagement profile of different engagement cohorts in the class and to know what characteristic are prominent among different cohorts. One can also view the details of each student in each cohort and assign specific tasks such as revisions and assessment retake. The *learning materials* interface show in Figure 4 provides stakeholders a way to access the learning materials students have used in the past including: textbooks, quiz questions, students' solutions and lecture slides. Figures 1 – 4 are from a real implementation of the Boll system currently deployed at a school in Japan.

4. Research Method

In this research, we use the Boll system (Ocheja, Flanagan, Ueda & Ogata, 2019) to connect the learning behaviour logs of students across two schools in Japan. We first setup the Boll system, connect it to the Learning Records Store (LRS) of the Junior High School (JHS) and assign a blockchain address to each student. The Boll system also keeps track of each student's ID at that school. This is then used to identify the records to be transferred when the student change school. When students in these schools move from the JHS 3 to High School (HS) 1 (a different school), we also transfer their past learning logs on the BookRoll system (Flanagan & Ogata, 2018) to their new school. The HS also has a similar setup of the Boll system with connections to the LRS.



Title	Course	Description	Type	Action
About math classes from the second half of 00	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	BOOK CHAPTER	OPEN
01 Realistic time calculation 1 Problem	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	QUIZ	OPEN
02 How many years has the universe been born? 1 problem	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	QUIZ	OPEN
03 Car stop distance 1 Problem	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	QUIZ	OPEN
04 Three Squares Theorem Bline Test Fix	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	QUIZ	OPEN
04 Train and Bicycle 1 Problem	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	QUIZ	OPEN
05 Final test for the previous term	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	EXAMINATION	OPEN
05 number S late mid-term test commentary	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	EXAMINATION	OPEN
06 Late mid-term test repair	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	EXAMINATION	OPEN
07 Airplane flight distance and time	2020 Mathematics [JHS 3] Group A	2020 Mathematics [JHS 3] Group A	BOOK CHAPTER	OPEN

Figure 4. Past Learning Materials Transfer across Schools.

For this study, we analyzed the learning behaviour logs of 109 students in JHS 3 Mathematics course in 2020 academic year who are currently in HS 1 and have enrolled in the HS 1 Mathematics course in 2021 academic year. Our analysis includes: engagement behaviour cohorts, temporal and spatial change in engagement and learning contents visualization. We measure engagement as a sum of different student behaviours categorized in to 5 dimensions: self-evaluation (S_e), cognitive behaviour (C_b), backtracking behaviour (B_b), time commitment (T_c) and content progress/completion (C_p). We define self-evaluation (S_e) as the students' ability to evaluate correctly their own solution to quiz questions. S_e is calculated as a fraction of the quiz answers from the student which were correct and rightly marked as correct by the student. Cognitive behaviour (C_b) is a measure of the students' cognitive action through cognitive indicators such as yellow and red markers added on learning materials through the BookRoll system (Akçapinar, Hasnine, Majumdar, Flanagan & Ogata, 2019). The backtracking behaviour (B_b) is an indication of how often students revisit concepts in order to improve their understanding or master such concepts. This is calculated as a weighted sum of total previous page visit actions divided by the total next page visit actions and the total previous page visit actions (Yang, Chen, & Ogata, 2021). Time commitment (T_c) is a measure of how often students study and it is calculated as the weighted sum of the total time, total number of content usage events and the total number of unique days students used the contents of the course. Content progress/completion (C_p) is a measure of how students advance towards completing the study materials. It is calculated as the weighted sum of total open and next page actions and total sum of long and short events. It is important to note that the parameters of each engagement metric were percentile rank of their actual values. Thus, student overall engagement is calculated as:

$$\text{Engagement} = S_e + C_b + B_b + T_c + C_p$$

In table 1, we show a summary description of the dataset for the JHS 3 Math course in 2020. The engagement metrics previously discussed were extracted from the dataset of the students who took the final exam and were graded. The engagement score was used to divide into quartile groups of 4 different engagement levels: Very High (≥ 75^{th} percentile), High (≥50^{th} percentile), Low (≥25^{th} percentile) and Very Low (< 25^{th} percentile) using percentile rank. We then proceeded with ensuring the data meet the assumptions of a one-way Analysis of Covariance (ANOVA) before performing conducting a test for a significant difference in the mean score for each engagement level. Finally, we developed 4 visualizations for teachers to view students' past engagement showing information such as: learner profile, group engagement, temporal and spatial engagement change and learning materials used.

Table 1. *Description of the Dataset*

	No. of Students	Total Logs	No. of Students graded
Group A	40	123,678	38
Group B	40	98,080	38
Group C	40	125,619	33

5. Results

Before carrying out an Analysis of Covariance (ANOVA) between the engagement levels and score, a Shapiro-Wilk test was conducted to determine the normality of the data. The result (0.99, $p > 0.05$) revealed that the score data across the different engagement levels followed a normal distribution. A further test for homogeneous variance using Levene's test indicated homogeneity of variances across the different engagement levels ($F (3,105) = 2.272$, $p > 0.05$). We then conducted a parametric one-way ANOVA to determine whether the mean scores of all engagement levels are different. The result ($F(3,105) = 3.783$, $p < 0.05$) indicated a significant difference in the mean scores for all engagement levels. A further post-hoc test using the Games-Howell test (due to unequal sample sizes) showed that the difference between very high and very low engagement levels is significant ($p < 0.05$) as presented in table 2. The implication of this result is that very low and very high engagement levels are indicative of the final performance of students and provide actionable insights for guiding future teaching and learning.

Table 2. *Post-Hoc Test (Games-Howell) Results of Scores between Engagement Levels (Mean Difference, Standard Error)*

	N	Score (μ)	SD	Very High	High	Low	Very Low
Very High	28	59.64	16.01	-	3.72 (4.12)	6.50 (4.02)	14.53 (5.06)*
High	27	55.93	14.51		-	2.78 (3.85)	10.82 (4.92)
Low	27	53.15	13.76			-	8.04 (4.84)
Very Low	27	45.11	21.07				-

Note. * $p < 0.05$.

6. Discussions

This work makes an important contribution of investigating and informing stakeholders the effect of students' prior engagement on their final scores at a different learning environment. Such information makes it possible for teachers to provide specific interventions at the start of a new class without having to wait to collect some data in the first few weeks. Although the results from our analysis only revealed a significant correlation between the scores and engagement of very high and very low engagement students, we propose this type of analysis to be performed when providing stakeholders with learning logs from a different learning environment.

In addition to engagement and final scores, this work provided access to resources such as the students' solution to examination questions and learning materials used. Access to this type of data give teachers additional information about the students' ability, and challenges with respect to the assessment questions. We acknowledge that in some cases, other contextual information may be required to correctly interpret the engagement measures extracted from the learning logs. Also, students may have received other scores different from the final score. It may be useful to consider how the students' engagement at intervals preceding other assessment affected their performance.

7. Conclusion

The transfer of learning logs and materials across different schools is an important requirement to solving the cold-start problem. This work presented metrics for defining engagement levels of students transitioning different learning environments. To validate the usefulness of the transferred data, we conducted some statistical tests which showed that the proposed engagement measures had a significant impact on students' final score. We also presented how teachers can access additional data from the students past learning data such as quiz answers by students, and the textbooks used as well as the previous teachers' lecture materials. Future work will be focused on validating the usefulness of the proposed visualizations with key stakeholders and the impact of our system.

Acknowledgements

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (B)20H01722, JSPS Grant-in-Aid for Scientific Research (S)16H06304 and NEDO Special Innovation Program on AI and Big Data 18102059-0.

References

Akçapinar, G., Hasnine, M. N. , Rwitajit, M., Flanagan, B., & Ogata, H., (2019). Exploring the Relationships between Students' Engagement and Academic Performance in the Digital Textbook System. 27th *International Conference on Computers in Education*, Taiwan.

Baker, R. S. (2019). Challenges for the future of educational data mining: The Baker learning analytics prizes. *Journal of Educational Data Mining*, *11*(1), 1-17.

Boekaerts, M., & Corno, L. (2005). Self- regulation in the classroom: A perspective on assessment and intervention. Applied Psychology, 54(2), 199-231.

Flanagan, B., & Ogata, H. (2018). Learning analytics platform in higher education in Japan. *Knowledge Management & E-Learning: An International Journal*, *10*(4), 469-484.

Hu, M., & Li, H. (2017). Student engagement in online learning: A review. In 2017 *International Symposium on Educational Technology*, 39-43. IEEE.

Li, L. Y., & Tsai, C. C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. *Computers & Education*, 114, 286-297.

Lu, O. H., Huang, J. C., Huang, A. Y., & Yang, S. J. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, 25(2), 220-234.

Majumdar, R., & Iyer, S. (2016). iSAT: a visual learning analytics tool for instructors. *Research and Practice in Technology Enhanced Learning*, 11(1), 1-22.

Martin, A. J., & Liem, G. A. D. (2010). Academic personal bests (PBs), engagement, and achievement: A cross-lagged panel analysis. *Learning and Individual Differences*, 20(3), 265-270.

Ocheja, P., Flanagan, B., & Ogata, H. (2018). Connecting decentralized learning records: a blockchain based learning analytics platform. In Proceedings of the 8th *international conference on Learning Analytics and Knowledge*, 265-269.

Ocheja, P., Flanagan, B., & Ogata, H. (2019a). Decentralized E-Learning Marketplace: Managing Authorship and Tracking Access to Learning Materials Using Blockchain. In *International Cognitive Cities Conference,* 526-535. Springer, Singapore.

Ocheja, P., Flanagan, B., Ueda, H., & Ogata, H. (2019). Managing lifelong learning records through blockchain. *Research and Practice in Technology Enhanced Learning*, 14(1), 1-19.

Vytasek, J. M., Patzak, A., & Winne, P. H. (2020). Analytics for student engagement. In *Machine learning paradigms*, 23-48. Springer, Cham.

Yang, C. C., Chen, I. Y., & Ogata, H. (2021). Toward Precision Education: Educational Data Mining and Learning Analytics for Identifying Students' Learning Patterns with Ebook Systems. *Educational Technology & Society*, 24(1), 152-163.