# Improving Learning Analytics and Student Performance through Connected Lifelong Learning on the Blockchain

**Ocheja Patrick, Flanagan Brendan, Lecailliez Louis and Ogata Hiroaki**
Kyoto University
ocheja.ileanwa.65s@st.kyoto-u.ac.jp; flanagan.brendanjohn.4n@kyoto-u.ac.jp; louis.lecailliez@outlook.fr; ogata.hiroaki.3e@kyoto-u.ac.jp

**ABSTRACT**: The use of learning analytics to enhance learning at different levels has continued to gain attention over the past few years. With learning activities taking place in different environments, systems and contexts, capturing and sharing these actions/outcomes continue to pose serious problems. In this paper, we examine how student learning interactions and events across various learning systems such as BookRoll can be accessed, transferred and protected across different learning environment. In facilitating the reusability of past learning interactions, we propose the transferability of user models derived from these learning logs. Finally, we discuss potential areas for advancing the field in connecting and using distributed learning logs for improving learning analytics and students' performance in general.

**Keywords**: lifelong learning, learning analytics, user model, privacy, learning record store, learning management system, blockchain, smart contract

## 1 INTRODUCTION

It is common to experience learning in different ways ranging from informal to formal contexts, and passive to active engagements. However, due to the ubiquitous nature of learning, it is difficult to capture and unify learner data from different environments as commonly experienced in other Big Data environments (Kadadi, Agrawal, Nyamful & Atiq, 2014). In this paper, we focus on discussing challenges to connecting lifelong learning data of students across different learning environments. Lifelong learning log is a journal that contains all of the learning activities carried out by a learner, and consists of a sparse multisource dataset of the learning actions of a learner. As learners change school, it is important to enable learning traceability by connecting learning experiences, and revisit consistent stakeholder concerns on data privacy and security. We demonstrate how lifelong learning can be achieved using blockchain technology and present some results from a live deployment in a K-12 environment.

### 1.1 Related Work

Lifelong learning is desirable and useful for learning analytics (Mouri & Ogata, 2015; Bakharia, Kitto, Pardo, Gašević & Dawson, 2016) as it provides ways to understand what a learner knows beyond current assessment and probe in detail the origin of difficulties or excellence. Due to the limited solutions that can facilitate data continuity across different learning environments, the act of combining data from multiple sources becomes a common alternative when lifelong learning data of learners is difficult to obtain or store.

Samuelsen, Chen, & Wasson (2019) in a review on multisource data for learning analytics identified the lack of tools or research work on meaningful data integration, storage and processing when combining multisource data. Kay & Kummerfeld (2019) proposed a conceptual model for evaluating how learning applications and data repositories can be used to realize Personal User Model for Lifelong, Life-wide Learners (PUMLs). PUMLs is proposed as a personal repository of raw data and interpretations that could be accessed by authorized programs. One question that comes up when implementing PUMLs is how to integrate with existing tools. Kay & Kummerfeld (2019) acknowledged this concern and mentioned interoperability as one of the requirements for realizing PUMLs. Another challenge with the PUMLs framework is how to manage privacy especially when user models are centrally stored and learners or their institutions may become less aware of how such models are been used.

To solve the limitations of centralizing lifelong learning, Ocheja, Flanagan & Ogata (2018) and Ocheja, Flanagan, Ueda & Ogata (2019) proposed a decentralized architecture where Learning Record Stores (LRS's) containing learner data are connected together through the blockchain and privacy permissions are managed using smart contracts. From the outcome of previous implementations, this paper discusses potential solutions to overcome the challenge of sharing and reusing learning experiences for analytics.

## 2 SOLUTIONS TO CURRENT PROBLEMS

As identified in the review of existing works on facilitating and enabling lifelong learning in the previous section, there is consistent concern with how to manage resulting data from learning tools, maintain privacy of learner information and facilitate learning analytics. In this work, we propose possible solutions to the issues of connecting distributed learning infrastructure, managing learning data, transferring user models and ensuring privacy.

### 2.1 Decentralization

As learning is ubiquitous in nature, and therefore realizing lifelong learning should not be the sole responsibility of one organization, but rather a function of places in which learning occurs. To capture activities of learners in different environments and institutes, Ocheja, Flanagan & Ogata (2018) proposed a decentralized learning analytics platform. A decentralized network makes it possible for multiple systems controlled by different actors to interact and transparently reach a consensus on protected resources. For example, to offer more privacy control on PUMLs (Kay & Kummerfeld, 2019), it becomes necessary to decentralize the user model such that when students change institution, they move with their model and update it at their new school or learning environment as new learning activities occur. For e-book learning logs, a decentralized access to lifelong learning makes it possible to obtain additional information outside the e-book context but impactful on learning outcome.

### 2.2 Tracing Learning

As learners move from one institute to another, it becomes necessary to know which institutes they have been to previously. One reason for such requirement is a case where a teacher needs to trace the root cause of a particular difficulty experienced by their student. In figure 1, *Bob's* teacher is faced with the task of detecting the gap in *Bob's* past learning in a prerequisite course. To detect this gap, *Bob's* teacher needs to know what topics in *Statistics* were covered at *Bob's* previous school and what *Bob's* performance and mastery was in each of these topics, which could be found in the learning logs from previous education institution. Connecting learning records on the blockchain provides an additional benefit of enabling traceability. This can be achieved using the nested transactions feature which is fundamental function of the design in blockchain where the current block contains a reference to the previous block.

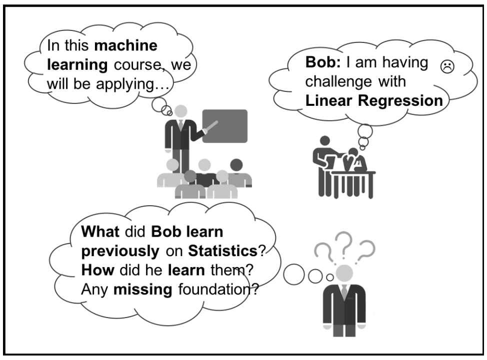


Figure 1: Tracing a learner's learning path.

## 2.3 Privacy

The lack of protection and control of private information by data owners exist as a result of the disconnection between different LRS's. An example of this problem can be seen when students move from one school to another and in the process are less aware of how their past learning data is being used. Although, learning analytics helps in improving the performance of learners (Okubo, Yamashita, Shimada & Ogata, 2017), the gains of learning analytics must be commensurate to respecting learner's privacy and associated rights (Rubel & Jones, 2016). While connecting learning logs across different systems and engendering transfer of these logs, it is necessary to prioritize learner's privacy: learners should be constantly aware and have control of their data. As proposed system connects all of a learner's data across institutes, it allows learners to still control their data at previous institutions through use requests even after they have ended their formal association with an education institution. This also enables the possible reuse of data for research purposes by providing a method of formally requesting access and use of a learner's data even after they have left an institution.

## 2.4 Shareable and Reusable Learner Models

Connecting and transferring learning data can be useful for creating learner models, however under-resourced institutions are often at a disadvantage because of data and computational limitations. By enabling sharing of models and data between institutions it may support the realization of more accurate learner models, and therefore it should be possible to allow learners to transfer these models across different institutes where they are enrolled. This will help to advance the learning personalization process and reduce the computational cost and technical requirements for recomputing a learner's model whenever they change school (Baker, 2019).

# 3 BOLL AS A TOOL FOR LIFELONG LEARNING AND ANALYTICS

We present the Blockchain of Learning Logs (BOLL) proposed in Ocheja et al. (2018 & 2019) and demonstrate how it enables lifelong learning and analytics with results from a live deployment at K-12 school. BOLL is a platform that enable learners to connect their lifelong learning events as verifiable and non-modifiable transactions on the blockchain. In a practical sense, learners can move from one learning institution to another and at the same time take all their past learning actions with them. We consider this solution a notable example as it provides answers to questions on privacy, decentralized access, transfer of records, tracing learning, and openness for integration. As BOLL is currently in active development, we present to the research community a framework that serves as a recipe for enabling lifelong learning and at the same time open for collaboration.

BOLL in one stretch, solve problems on data privacy, trust between stakeholders and third parties and the overhead of manually transferring and collecting learner data across different institutions and systems. The use of a decentralized technology effectively ensures that no single party can abuse the interest of others without having at least the consent of 51% or more members on the network (Nakamoto, 2008). In this light, the questions we now ask revolve around how to on-board various institutions, facilitate a seamless transition and advance the field along this path.

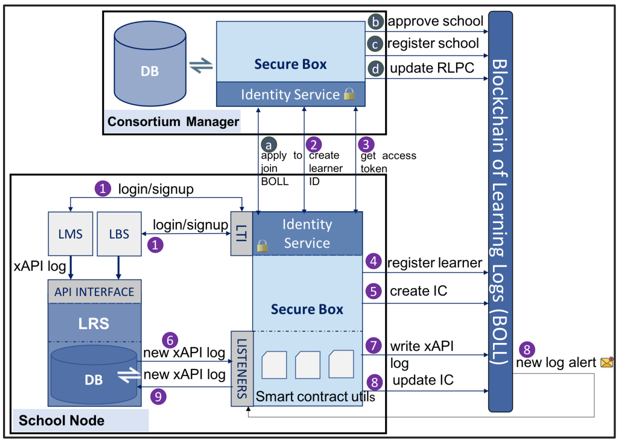


Figure 2: BOLL enrolment procedure.

## 3.1 Onboarding Requirements

Learning organizations can join BOLL and become stewards to safeguarding the learning records of learners across all institutes on the network. These learning organizations are regulated through a consortium which ensure members are verified and accountable. In addition to existing learning technologies, each institute is required to have an Ethereum blockchain node and a SecureBox: an open source by application project containing a set of utilities for interacting with the blockchain.

As for learners, they can join BOLL through their host institution. Because each learning institute has its own authentication system, BOLL allows each institute to connect to their authentication system using a Learning Tools Interoperability (LTI) module. This procedure is mandatory at least for the first-time a learner accesses BOLL and processes 1–5 in figure 2 are carried out. Subsequent access to BOLL is authorized through OAuth2 provided by the consortium manager. In a case where learners are underage and require parental consent (e.g. K-12 learners), BOLL provide settings for automatically generating the accounts for such learners when their parent or guardian opts-in to logging their ward's learning actions on BOLL. In figure 3, we show a distribution of some on-boarding processes from a live deployment of BOLL at a K-12. Creating smart contract for records indexing, registering user and pairing learners to institutes are done automatically for the K-12 learners.

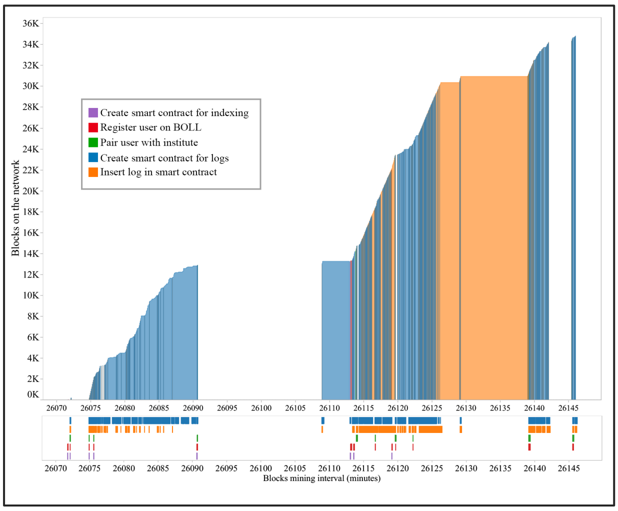


Figure 3: Snapshot of BOLL network deployed at K-12.

### 3.2 Privacy: Access and Authorization

BOLL facilitates learner privacy by using smart contracts as proposed by Ocheja et al. (2018) and implemented by Ocheja et al (2019). BOLL groups learning data according to the action verb that denotes the learning action performed. BOLL then assigns a specific smart contract to each type of learning action. A learner may grant a *read, write and/or admin* permissions to another party. A read permission allows the party to view the learner's records. A write permission allows the party to write learning logs on behalf of the learner. Only approved learning institutes can write these data for learners. An admin permission gives the party access to read and write learning actions for the learner and also allows the party to give other parties similar permissions on the learner's data. In the current implementation of BOLL, permissions are based on action verbs of the IMS Caliper (IMS Caliper, 2013) and/or the Experience API (xAPI) (Advanced Distributed Learning, 2016) specifications. BOLL recommends the use of open learning logs standards such as the xAPI and IMS Caliper.

### 3.3 Creating and Viewing Learning Logs

The task of creating learning logs on BOLL is done in the background while learners interact with learning tools connected to BOLL. As shown in steps 6 and 9 on figure 2, listeners are configured to pick up new learning events and write them to the blockchain or another institute's LRS. To issue testimonials such as certificates, recommendation letters and other documents, BOLL provides a view for staff members of an institution to create and issue such documents. All issued testimonials contain the cryptographic signature of the issuer. Most of the transactions from the live deployment of BOLL result from creating and inserting learning events as shown in figure 3. BOLL also provides an interface for learners to view their learning logs at different institutions.

### 3.4 Connect and Transfer Learning Logs

Connection of learning data is enabled on BOLL by writing each learning action on the blockchain. When a record is written on the blockchain for a learner, all institutes that have permission to read their learning records are notified through global events emitted on the blockchain. These institutes can then request for a copy of the new learning log from the originating institute's LRS. To transfer past learning records to a newly approved institute, the SecureBox contain functions that can query past LRS's where a learner has previously schooled to get their past data. These functions are automatically triggered by a *permission-granted* event immediately the learner grants access to the new institute.

### 3.5 Making Sense of Lifelong Learning

BOLL provide some useful features for understanding a learner's past learning experience. For example, a teacher can ask their student for permission to view the student's lifelong learning on BOLL. When the permission is granted, the teacher can view the courses their student has previously taken. In order to understand how their past learning could be related or useful to the teacher's course, the teacher can pick one of several models to run on the student's past logs. This view is shown in figure 4. The teacher first selects which course on the LMS or learning tool they want to view student's lifelong learning. After selecting the student, the teacher can then select which of the learner's past schools to get the learner's lifelong learning. Because the past learning events can be of different types, the teacher can specify which type of learning actions to include in the retrieved data. The final step is to decide which learning analytics model to apply on the retrieved data. Here, we identify a potential for BOLL to integrate with other learning analytics dashboard and visualization tools such as that proposed by Majumdar, Akçapınar, Akçapınar, Ogata & Flanagan (2019).

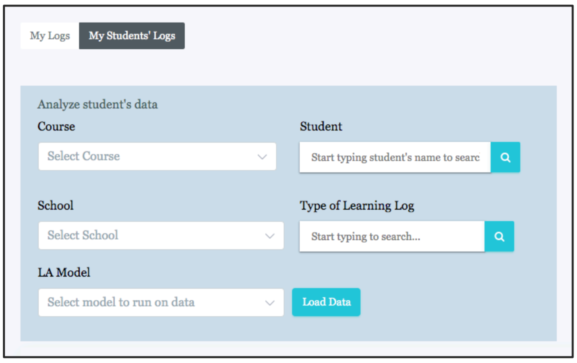


Figure 4: BOLL interface for Analyzing Lifelong Learning.

Another example of how BOLL makes sense of lifelong learning is in helping teachers understand a learner's knowledge map. For example, a teacher may want to know what a learner knows in relation to a particular subject. A learner's knowledge map can then be constructed with their lifelong learning on BOLL. Thus, BOLL can serve as a data backend for knowledge map analysis tools like the tool proposed by Flanagan, Majumdar, Akçapınar, Wang & Ogata (2019). Flanagan et. al (2019) proposed a knowledge map creation platform capable of computing a learner's knowledge states over time, grouping knowledge states, intra-cohort comparison and computing relationship between various knowledge states. Such visualizations provide a quick and concise view of what a learner knows and what they may find easy or difficult to grasp.

### 3.6 Metadata: Public and Private

When transactions representing learning actions of students are written on the blockchain, all participants on the network get notified. But because the BOLL network does not store the actual learning actions on the blockchain, it is not possible to see what learning actions learners performed. However, when some metadata such as test scores or grades are stored directly on the blockchain, the decision on whether to make such contents public or private becomes a concern for stakeholders. One way to ensure that scores which are expected to be kept private remain so is that such scores should not be included as metadata. Another alternative could be to redact such scores by using some non-publicly communicated offsets or encryption before storing them on the blockchain. For a consortium, it is more appropriate to provide standards to guide all members but some institutions may decide on what works best for them. In this case, any choice made should be communicated to other parties that want to make sense of such data.

## 4 FUTURE IMPLICATIONS FOR THE FIELD

As the field of lifelong learning and analytics continues to evolve, the following important topics remain and require investigation in greater detail. By extending the framework proposed in this paper, it may be possible to overcome part of these problems, and it would require more collaboration with stakeholders and the wider research community.

*Transferring and connecting learner/learning models*: When a learner's learning data is used to construct a model, it is useful to enable the learner to keep a record of such model and be able to grant their subsequent learning institutions access to such models. One of the arguments in favour of this feature is a case where replicating the same model might be unachievable especially when different institutions have varying access to different learning analytics technologies. Thus, it is necessary to understand the requirements for sharing learner/learning models and enable such possibilities when connecting lifelong learning.

*Off-the-shelf learning analytics for connected lifelong learning*: Connecting lifelong learning becomes useful if teachers and/or students can obtain meaningful feedback from the connected data. To help them gain insights from these connected data, it is necessary to provide tools for analytics and visualization. While there are existing efforts in providing learning analytics dashboards, future concerns should be focused on decentralizing learning analytics dashboards such that they can be easily integrated with data coming from multiple institutions. Also, because the connected lifelong learning data could be of different types (xAPI standard, IMS Caliper standard, etc.) it is also necessary to develop a framework to unify these different standards.

*Integrating different learning logs standards*: As lifelong learning involves learning actions of a learner at different schools where different learning technologies are used, it is necessary to ensure that such learning actions are compatible or can be merged. A common way of ensuring compatibility is the adoption of a standard such as xAPI for expression learning actions. However, it is possible that some institutions may prefer a set of learning events standards over another. Therefore, we recommend the development of a framework to unify learning logs resulting from different standards. While this task might be trivial in a case where learning actions are expressed in well-defined standards, this task becomes more difficult when no proper documentation on the format of learning actions are provided. A community driven approach could be to maintain a public repository that list all available standards for learning actions stored on BOLL. This makes it possible for learning analytics tools to conveniently interpret connected lifelong learning regardless of the variation in data format.

*Demise of an institution*: As only a hash of the learning log and its location is recorded on BOLL, there is a possibility of a learning log outliving its host institution. To illustrate this, a student might graduate from an institution and 10 years later, that institution ceases to exist. In a case where all computing facilities such as the LRS of that institution is shutdown, then the learning logs whose references are held on the blockchain cannot be retrieved anymore. To solve this problem, it will be helpful to have a learning blockchain with third parties who can offer data backup services.

## 5 CONCLUSION AND FUTURE WORK

In this work, we assessed various contributions in the field of learning analytics towards the realization of connected lifelong learning and use of multisource data to improve learning. A review of previous research shows that being able to access lifelong learning logs is useful for learning analytics. But this is still lacking in the field especially due to multiplicity of learning tools, security and privacy concerns, and the administrative burden of using and managing existing solutions.

To solve the problem of disconnected lifelong learning and challenges with adoption of potential solutions to existing problems, this paper presented new pivots that can be used to advance the field of learning analytics towards the realization of connected lifelong learning. One of such pivots is the decentralization of learning analytics and tools such that learning organizations can operate as smaller units in a large network with lesser administrative burden when deciding rules on data collection, access, usage and transfer across institutes. The BOLL network is presented as a solution to an existing challenge of ensuring trust among institutions and the difficulty in deploying learning analytics on a wider scope when policies are continuously evolving. Future work will be focused on providing answers to challenges presented as implications for the field including transfer of learning models, integrating off-the-shelf learning analytics, unifying learning logs standards and conducting user studies to measure the impact of the proposed BOLL system.

## ACKNOWLEDGEMENT

This research was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (S) Grant Number 16H06304.

## REFERENCES

Advanced Distributed Learning. 2016. Experience API (xAPI) Specification. http://github.com/adlnet/xAPI-Spec. Accessed September 10, 2019.

Baker, R. S. (2019). Challenges for the Future of Educational Data Mining: The Baker Learning Analytics Prizes. *JEDM| Journal of Educational Data Mining*, 11(1), 1-17.

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016). Recipe for success: lessons learnt from using xAPI within the connected learning analytics toolkit. *In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 378-382). ACM.

Flanagan, B., Majumdar, R., Akçapınar, G., Wang, J., & Ogata, H. (2019). Knowledge Map Creation for Modeling Learning Behaviors in Digital Learning Environments. *In Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge* (DC@ LAK19).

IMS Global Learning Consortium. (2013). Learning measurement for analytics whitepaper. *IMS Learning Analytics White paper*, 1-12.

Kadadi, A., Agrawal, R., Nyamful, C., & Atiq, R. (2014). Challenges of data integration and interoperability in big data. *In 2014 IEEE international conference on big data* (pp. 38-40). IEEE.

Kay, J., & Kummerfeld, B. (2019). From data to personal user models for life‐long, life‐wide learners. British Journal of Educational Technology, 50(6), 2871-2884.

Majumdar, R., Akçapınar, A., Akçapınar, G., Ogata, H., & Flanagan, B. (2019). LAViEW: Learning Analytics Dashboard Towards Evidence-based Education.

Mouri, K., & Ogata, H. (2015). Ubiquitous learning analytics in the real-world language learning. *Smart Learning Environments,* 2(1), 15.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

Ocheja, P., Flanagan, B., & Ogata, H. (2018). Connecting decentralized learning records: a blockchain based learning analytics platform. *In Proceedings of the Eighth International Conference on Learning Analytics & Knowledge* (pp. 265-269). ACM.

Ocheja, P., Flanagan, B., Ueda, H., & Ogata, H. (2019). Managing Lifelong Learning Records Through Blockchain. *Research and Practice in Technology Enhanced Learning*, 14(1), 4.

Okubo, F., Yamashita, T., Shimada, A., & Ogata, H. (2017). A neural network approach for students' performance prediction. In *International Conference on Learning Analytics and Knowledge* (pp. 598-599).

Rubel, A., & Jones, K. M. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, 32(2), 143-159.

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics review of literature. *Research and Practice in Technology Enhanced Learning*, 14(1), 11.