Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

OKLM: Open Knowledge and Learner Model Using Educational Big Data

Kensuke TAKII[a*], Changhao LIANG[b] & Hiroaki OGATA[b]
[a]*Graduate School of Informatics, Kyoto University, Japan*
[b]*Academic Center for Computing and Media Studies, Kyoto University, Japan*
*kensuke.takii96@gmail.com

Abstract: This study proposes the Open Knowledge and Learner Model (OKLM), a novel framework that integrates Learning Analytics (LA) with Digital Twin (DT) technology to model learners' knowledge, internal states, and environments. The OKLM DT framework addresses the limitations of traditional LA systems by enabling accurate estimation of knowledge states and personalized learning strategies. We developed a conceptual framework for the learner DT using LA, verified the accuracy of the OKLM-based DT model, and applied it to a learning support system. Initial experiments in an English literature recommendation system showed that, while the recommendations did not significantly enhance learners' motivation, they were well-received and positively correlated with increased engagement among highly motivated learners. A subsequent study involving Intensive Reading (IR) support for EFL learners further validated the model's effectiveness. Additionally, experiments targeting educators demonstrated that the OKLM's visualization tools were valuable for understanding learner characteristics and tailoring teaching materials. These findings suggest that OKLM can enhance the versatility and accuracy of learner models across various educational contexts, offering a significant advancement in the field of LA.

Keywords: Open Knowledge and Learner Model, learner digital twin, learning analytics, data-driven system, model validation, visualization

1. Introduction

Teaching and learning aim to expand knowledge and foster growth. Effective education requires understanding learners' internal states (e.g., knowledge, psychology) and external factors (e.g., content, environment). Traditional educational settings, with few teachers and many students, make it difficult to assess individual needs and provide optimal support. The complexity of learning further complicates this. Learning Management Systems (LMS) and Learning Analytics (LA) (Ferguson, 2012) have emerged to address these challenges by analyzing learner activities. However, current LA systems struggle to accurately estimate knowledge states and determine efficient learning strategies, limiting personalized support.

In response, we propose the Open Knowledge and Learner Model (OKLM), a framework combining LA and Digital Twin (DT) technology to model learners' knowledge, internal states, and environments. The OKLM DT simulates these elements in a virtual space, using LA to collect, analyze, and model data. Our research has three goals:

1. Develop a conceptual framework for learner DT using LA.
2. Verify the accuracy of the OKLM-based DT model.
3. Apply OKLM to a learning support system and assess its effectiveness.

To achieve these goals, we adapted a DT model for learner modeling and implemented the OKLM DT. We tested the model's accuracy against actual learning data. Finally, we will implement a learning support system using OKLM DT, evaluate its impact, and discuss expected contributions and future work.

2. Related Works

2.1 The Concept of Knowledge and Learner Models

In supporting human learning and education with information and communication technology, constructing a knowledge model (KM) is crucial for understanding the learner's knowledge state. While manual KM construction is labor-intensive and requires expertise (Khadir et al., 2021), recent research has advanced automatic KM construction using computers to extract knowledge from learning materials (Chen et al., 2018). The learner model (LM) represents the learner's knowledge state and, together with KM, is vital for educational support systems. In Open Learner Model (OLM) research (Bull, 2020), allowing learners to access and modify their LMs improves accuracy and learning outcomes. However, traditional OLMs are often tied to specific knowledge areas, limiting their accuracy and versatility. The OKLM aims to enhance versatility by integrating LA technology into OLMs.

2.2 Learner's Digital Twin

DT is a virtual replica of a physical product or system, reflecting its state using real-time data (Grieves, 2014). Tao et al. (2018) proposed a five-tuple DT model including physical and virtual entities, services, DT data, and connections among them. The concept extends to humans as the Human Digital Twin (HDT) (Okegbile et al., 2022), primarily applied in healthcare. While there are proposals to use HDT in learning contexts, such as Furini et al. (2022), practical demonstrations are lacking. Incorporating not only the learner's knowledge state but also their internal state and environment into the model is crucial for achieving adaptive learning (Weber, 2012).

3. Research Framework: Infrastructure of OKLM

Figure 1 shows a concept of OKLM. OKLM has following five remarkable features: (1) Openness of the knowledge and learner models, (2) Visualization of the knowledge and learner models, (3) Future predictions of learners' states, (4) Analysis of past grades or events, and (5) Providing media for external people or systems to apply OKLM to learning support.

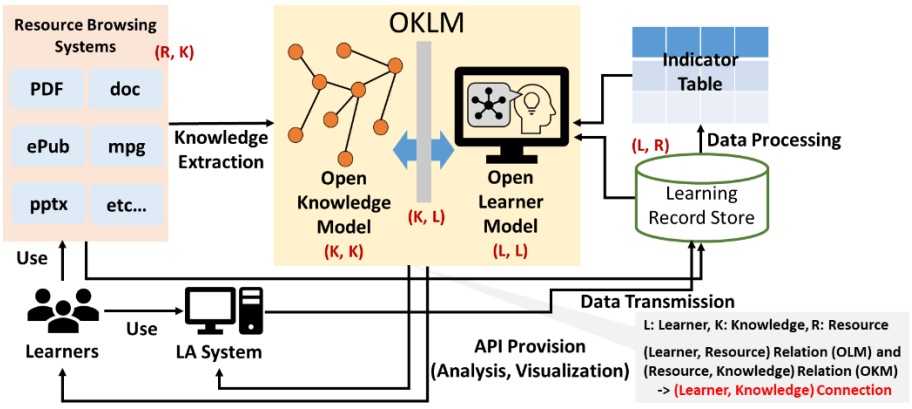

Figure 1. OKLM Framework.

OKLM consists of a knowledge model and a learner model. To realize the knowledge model, A system that automatically generates it—a graph structure of knowledge items—from learning materials and links it with learning logs is necessary. OKLM can use an existing method (Flanagan et al., 2019) to create this map by analyzing the co-occurrence of nouns in instructional materials, with words as nodes and co-occurrences as branches. To model learners' states and environments, learning logs should be connected to the knowledge model.

In the developed system, each node in the map stores information about the learning logs related to that knowledge item.

OKLM offers various possibilities for learning support. For instance, visualizing the knowledge map can help students understand the knowledge targeted by their activities and discover new knowledge through network analysis. By analyzing learner and material characteristics, and using statistical methods, OKLM can support peer learning, recommend materials, and predict grades.

4. Contributions to the Fields of LM and LA Research

The novelty of OKLM lies in its versatility. Traditional learner models are tied to specific learning support systems, limiting their ability to capture diverse learning behaviors and leading to inaccuracies. OKLM, supported by LA, allows any knowledge-acquisition behavior to inform model building, leveraging the diversity in collected data. The separation of the learning support system from the knowledge/learner model enables the system to handle knowledge across various fields and apply the model in multiple learning contexts. This versatility is expected to refine training data and advance existing LA techniques.

5. Proposed Research Methodology

To achieve the proposed research goals, we are planning to conduct experiments in various knowledge domains and learning contexts.

5.1 Learning Support with a Behavioral LM (Before OKLM)

To establish a baseline for comparison with OKLM, we conducted an experiment using a learning support system without a knowledge or learner model, in the context of Extensive Reading (ER) support for Japanese EFL learners (Takii et al., 2024). Participants used the e-book reader BookRoll (Flanagan & Ogata, 2018) within the LA platform LEAF (Ogata et al., 2023). The system recommends books based on reading logs, vocabulary, and difficulty but does not incorporate deeper knowledge analysis like OKLM.

We evaluated which learner characteristics are receptive to recommendations using data on system access, clicks on recommended book titles, and correlations between these activities and learners' impressions, measured by the Technology Acceptance Model (TAM) (Park et al., 2012). The results showed that while the recommendations did not impact learners' English skills or motivation, they were well-received, with a strong positive relationship between their use, acceptance, and motivation to learn. This study suggests that, although the system did not generally increase motivation, it can further enhance the motivation of highly motivated learners.

5.2 Learning Support with OKLM

In the subsequent study, we will validate the learning effects of OKLM-based systems through two use-cases: an English academic literature recommendation system and a system supporting Intensive Reading (IR) for EFL learners. The first use-case involves undergraduate students in an English literature class using the recommendation system, which suggested papers based on students' reading activities and vocabulary analysis. The second use-case, which will be conducted in Japanese junior-high and high schools, involves students reading and translating English sentences. The system collected data on students' answers, self-evaluations, and knowledge of English grammar to verify the accuracy of the learner models.

Regarding the first use case, all the processes of the experiment have ended, and we will conduct data analysis for further discussion. The second experiment is now being conducted. After the experiment, we will analyze the data and write a paper.

5.3 Validation and Visualization of OKLM

To investigate the OKLM's unique structure that is not found in traditional LMs, we need to validate the accuracy of learners' representations in the OKLM. Our research includes an investigation of the relationship between learning outcomes and features in the OKLM data.

OKLM also includes a visualization interface for learner models, displayed as a network of knowledge items and their relationships. We conducted an experiment to assess the impact of this visualization, targeting a teacher in an undergraduate course, and evaluated her impressions of the OKLM visualization. Targeting teachers, an experiment examining how helpful OKLM is in understanding learners' characteristics will be planned. Targeting students, we will measure the learning effect of examining the OKLM visualization and what kind of information they can obtain. The experiment targeting a teacher in an undergraduate course showed that she had a good impression of OKLM, finding it useful for understanding the characteristics of learners and the composition of teaching materials.

Acknowledgements

This work was supported by CSTI SIP Program Grant Number JPJ012347, NEDO JPNP20006, and JST Grant Number JPMJFS2123.

References

Bull, S. (2020). There are open learner models about!. *IEEE Transactions on Learning Technologies*, *13*(2), 425-448.

Chen, P., Lu, Y., Zheng, V. W., Chen, X., & Yang, B. (2018). Knowedu: A system to construct knowledge graph for education. *Ieee Access*, *6*, 31553-31563.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5-6), 304-317.

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 (LAK'19)* (pp. 428-436). Society for Learning Analytics Research (SoLAR).

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

Furini, M., Gaggi, O., Mirri, S., Montangero, M., Pelle, E., Poggi, F., & Prandi, C. (2022). Digital twins and artificial intelligence: As pillars of personalized learning models. *Communications of the ACM*, *65*(4), 98-104.

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper*, *1*(2014), 1-7.

Khadir, A. C., Aliane, H., & Guessoum, A. (2021). Ontology learning: Grand tour and challenges. *Computer Science Review*, *39*, 100339.

Ogata, H., Majumdar, R., & Flanagan, B. (2023). Learning and evidence analytics framework bridges research and practice for educational data science. *Communications of the ACM*, *66*(7), 72-74.

Okegbile, S. D., Cai, J., Niyato, D., & Yi, C. (2022). Human digital twin for personalized healthcare: Vision, architecture and future directions. *IEEE network*, *37*(2), 262-269.

Park, S. Y., Nam, M. W., & Cha, S. B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British journal of educational technology*, *43*(4), 592-605.

Takii, K., Flanagan, B., Li, H., Yang, Y., Koike, K., & Ogata, H. (2024). Explainable eBook Recommendation for Extensive Reading in K-12 EFL Learning. *Research and Practice in Technology Enhances Learning* (in press).

Weber, G. (1999). Adaptive learning systems in the World Wide Web. In *UM99 User Modeling: Proceedings of the Seventh International Conference* (pp. 371-377). Springer Vienna.