Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education.
Asia-Pacific Society for Computers in Education

Extracting Students' Self-Regulation Strategies in an Online Extensive Reading Environment using the Experience API (xAPI)

Chia-Yu HSU [a*], Izumi HORIKOSHI[b], Huiyong LI[b], Rwitajit MAJUMDAR[b] & Hiroaki OGATA[b]
[a]*Graduate School of Informatics, Kyoto University, Japan*
[b]*Academic Center for Computing and Media Studies, Kyoto University, Japan*
*hsu.chiayu.25t@st.kyoto-u.ac.jp

Abstract: Extensive Reading (ER) activity is useful in language learning where learners pick any reading materials in target language by themselves and continue reading. In this study, we aim to understand students' self-regulation behaviors and strategies in ER so that facilitators can give learning strategy-based instruction. For this purpose, we first explored a potential structure of a learners' model that can highlight their learning strategies and extracted from multi-system interaction logs recorded in Experience API (xAPI) format. To demonstrate, we collected data from 120 students across 3 months when they did ER activities. We analyzed and extracted the self-regulation strategies of the learners and found 2 groups by applying K-means cluster analysis. The results inform dashboard design and instructional support based on the visualized attributes of the cluster members. This study contributes towards using interoperability standards to record learner's online reading behaviors and demonstrate how teaching and learning activities can be supported by xAPI when such experiences are distributed across various learning tools.

Keywords: xAPI, Learning Analytics, Extensive Reading, Learning Strategies

1. Introduction

Considering the effect of improving students' reading efficiency, Extensive Reading (ER) is regarded as an important component in reading skills (Mustapa, Rahman, Ghani, Saad, & Mohamed, 2018; Takii, Flanagan, & Ogata, 2021). In recent years, the advance of ICT technology has enabled a variety of support for the learning from higher education to K12 (Lin et al., 2016). In terms of the selection of ER materials, some personalized e-learning recommender systems can adapt to learners' different interests and levels, such as the system proposed by Takii et al. (2021). An online tool, M-Reader, proposed by Robb and Kano (2013) allowed students to take a short quiz after finishing a book and reduced teachers' workload on keeping track of their students' reading. To support self-regulated learning, the GOAL system, proposed by Li, Majumdar, Chen and Ogata (2021), scaffolds students to schedule their learning activities and enables their data informed decision making.

The researchers suggested that to get students to read more, it may be necessary to create monitoring systems that are built to facilitate the implementation of ER by teachers (Chang & Renandya, 2017; Renandya & Jacobs, 2002; Mustapa et al., 2018). Frequent questions are often raised about how to integrate student data from heterogeneous learning contexts with different software components, and what should be tracked to obtain useful information and avoid noisy data (Carrillo, Lavoué, & Prié, 2016). Tin Can API, also known as Experience API or xAPI, has a simple vocabulary able to capture and share the activity flow of almost any tool, which makes it a potential mechanism to manage and visualize the data of learning experience, leading to more adaptive and personalized learning services (Lin et al., 2016).

In terms of teachers' monitoring for ER activity, it is crucial to understand students' self-regulation behaviors and strategies in ER so that teachers can give learning strategy-based instruction. Hence, we find it necessary for a potential structure that can help describe learners' behaviors and extract their self-regulation strategies from logs recorded in xAPI considering that ER activity can be supported by different technologies. 2 research questions are: (1) How can students' learning behaviors and strategies be extracted from logs recorded in xAPI? (2) What are the students' self-regulation behaviors and strategies in ER?

2. Foundation

2.1 Experience API (xAPI) to Extract Learning Strategies

Vast quantities of data on student's online activity are recommended as observable indicators, supporting valid inferences about a learner's learning if taken advantage of. Different Learning Analytics (LA) methods can be used to help make meaningful sense from the data of massive volume and high rate of velocity (Viberg, Khalil, & Baars, 2020; Raga Jr, Raga, & Cariño, 2018). Saint, Gašević and Pardo (2018) used process mining techniques to identify strategic and tactical learner behaviors and found that certain temporal activity traits relate to performance in the summative assessments attached to the course, mediated by strategy type. Carlson, Genin, Rau and Scheines (2013) used an expectation-maximization approach to extract the error-making and hint-seeking behaviors of each student to characterize their learning strategy. However, many of these methods are limited in their ability to describe and interpret differences between extracted latent representations at varying levels of granularity, such as underlying data of student actions and behavior.

xAPI is a specification for learning technology that allows recording a wide range of experiences that a user can have in different technologies or tools (Lin et al., 2016). An event is captured as a Statement. The main elements in a Statement are the Actor, the Verb (action) and the Object, which forms the simplest event case: "[actor] [verb] [object]" (Manso-Vázquez, Caeiro-Rodríguez, & Llamas-Nistal, 2018; Carrillo et al., 2016). The limited but expandable vocabulary makes it easy to record new actions, and hence makes the creation of specific domain profiles possible (Manso-Vázquez et al., 2018; Serrano-Laguna et al., 2017). The concept of application profile was defined as "a way to declare which elements from which namespaces are used in a particular application or project". xAPI-SRL, for example, provides a standard way to record all the actions and activities typical in self-regulated learning. Different use cases covered in an xAPI profile are usually formed by several recipes based on the development of Communities of Practice. In xAPI-SRL, the large number and variety of strategies results in different recipes into which the identified actions are divided. (Manso-Vázquez et al., 2018). That is, an xAPI profile can give structure to a whole learning activity, describing learning strategies as well as the study tactics that compose them.

2.2 Extensive Reading in the Learning Evidence Analytics Framework (LEAF)

In this study, we investigated students' behaviors of extensive reading in a Japanese junior high school, in which the learning activities are supported with various tools in the Learning Evidence Analytics Framework (LEAF). Figure 1 shows the learning tools composed of the framework along with the affordances of each tool. Examples of learner actions and the trace data generated are also given.

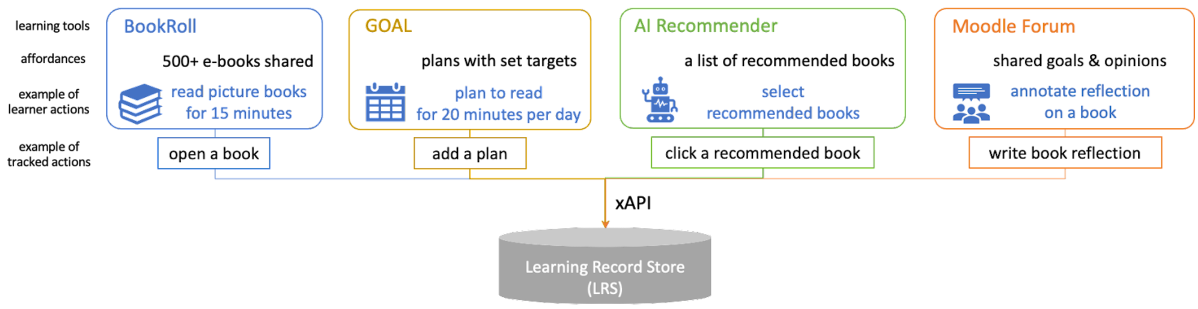

Figure 1. Extensive Reading in LEAF.

In the ebook reader, BookRoll (Ogata et al., 2015), students could read more than 500 English picture books with different levels. The self-directed learning scaffolding system, Goal-oriented Active Learning System (GOAL), supports students to make reading plans to achieve targets set by themselves. To help students find a suitable book, the material recommending system, AI Recommender, generates recommendations based on students' previous learning activities. With the Learning Management System (LMS), Moodle, teachers launched a forum in which the students could share their opinions with each other every month. The following describes an example action for which the xAPI object is generated in those tools. "open a book" is generated when students read picture books in BookRoll. Plans made in the GOAL system are recorded as "add a plan". When a book is selected in AI Recommender, students are recorded to "click a recommended book". The reflection shared in the Moodle Forum generates the action: "write book reflection". Trace data of students' actions are logged in the Learning Record Store (LRS).

3. Extracting Self-regulation Extensive Reading strategies

3.1 Aggregating the Actions

To extract the learning strategies implied by the actions, we aggregated the actions into a common xAPI statement structure with the xAPI-SRL profile (Manso-Vázquez et al., 2018), introduced in Section 2. Among 11 proposed recipes, we identified 6 recipes according to the functionality of the systems where the actions were taken. Table 1 summarizes the actions in the systems in LEAF as well as the verbs and activity types of objects which were used on the recommendation of different recipes. Different recipes can be used to describe different system functionality. For instance, the 3 recipes used for GOAL reflect the system's functionality of supporting learners' task and goal management, including planning their goals, and controlling the time and flow to complete tasks and achieve goals. In each of the recipes, different statements are used to describe actions with recommended vocabulary of verbs and activity types of objects for the actions. In the GOAL system, its functionality of task and goal management supports learners to add a plan. The recipe reflecting this function controls that the verb, created, is used to indicate the activity type of object, task. Hence, in the common xAPI statement structure, we defined the action of adding a plan as follows: the verb, created, should be used to indicate a plan added by a learner, and the plan (object) should be indicated as the activity type, task.

Table 1. *The Recipes Used to Aggregate the Actions in LEAF*

Recipe	Task and goal management			Planning	Time and flow control		View control	Annotation	Curation
SRER[1] Actions	add a plan	edit a plan	delete a plan	select WPM[2] as target	estimate reading time as target	suspend a plan	open a book	write book reflection	click a recommended book
Verbs	created	edited	deleted	selected	estimated-duration	suspended	read	annotated	selected
Activity Types	task	task	task	goal	goal	task	book	note	book
System	GOAL						BookRoll	Moodle Forum	AI Recommender

[1]SRER: self-regulated Extensive Reading
[2]WPM: words per minute

3.2 Extracting the Strategies

When developing the profile, Manso-Vázquez et al. (2018) divided complex composite strategies into simple strategies and identified the actions and steps learners should follow to use the strategy. They defined the information and data resulting from the performance of the strategy in the software, including the actors, action, object of the action, results, context, time, and so on. On the basis of the verbs and activity types of objects, we corresponded the actions in LEAF to 7 simple strategies as summarized in Table 2. For example, we defined the action of adding a plan as: a task is created. This action involves the simple strategy: Task definition/identification. Based on the categorization of

Manso-Vázquez et al. (2018), we also corresponded the strategy to the composite strategy it belongs to: Metacognitive strategies.

Table 2. *The Categorization of the Strategies Corresponding to the Actions in LEAF*

composite strategy	simple strategy	actions in LEAF
Metacognitive strategies	Task definition/identification	add/edit/delete a plan
	Time estimation	estimate reading time as target
	Time use evaluation	suspend a plan
Cognitive strategies	Reading speed	select WPM* as target
	Content/information review	open a book
	Annotation (note taking, highlighting)	write book reflection
	Selection of information sources	click a recommended book

*WPM: words per minute

4. Implementation of the Extraction from Multi-system Trace Data

4.1 Methods

The data were collected from a junior high school in Japan. 120 seventh-graders participated in ER activity from April to June in 2021. We aggregated the action logs into the xAPI statement structure described in Table 1. We then extracted the learning strategies from the actions using the categorization in Table 2. The strategies are represented by the total numbers of the corresponding actions taken. For example, we counted how many times a student added/edited/deleted a plan respectively. Since these actions involve the simple strategy: Task definition/identification, we considered the occurrence where the students had this strategy by summing up the times s/he added/edited/deleted a plan during the period. Based on the numbers of the simple strategies extracted, we used K-means Cluster Analysis to cluster the students. We identified 2 as the optimal number of clusters with the Average Silhouette Method.

4.2 Results

We could label the two clusters which emerged in K-means Cluster Analysis as: Metacognitive & Cognitive Learners (n=11) and Metacognitive Learners (n=109). Table 3 summarizes the average counts of each simple strategy in the 2 clusters.

Table 3. *The Statistics and Differences of Strategies in the 2 Clusters*

	Metacognitive & Cognitive Learners		Metacognitive Learners	
	M	SD	M	SD
Task definition/identification	3.73	0.90	4.60	4.04
Reading speed	1.64	0.67	1.72	0.97
Time estimation	1.64	0.50	1.61	1.05
Time use evaluation	0.00	0.00	0.26	1.88
Content/information review ***	39.10	12.89	5.90	7.63
Annotation (note taking, highlighting) ***	13.91	8.63	0.61	1.45
Selection of information sources	3.55	8.41	0.07	0.30

*** $p < .001$

The clusters indicate the composite strategies derived from the simple strategies extracted from the actions according to categorization in Table 2. In the 3-month ER activity, Metacognitive Learners tended to focus on organizing their reading plans but didn't spend much time on reading books and sharing their reflection (where the average counts of Metacognitive strategies are higher than those of Cognitive strategies). On the other hand, Metacognitive & Cognitive Learners made reading plans as often as Metacognitive Learners did (where their average counts of Metacognitive strategies are similar to those of Metacognitive Learners), and they also engaged in reading books either themselves or on the recommendation of AI Recommender and sharing their opinion in the Moodle forum (where the average counts of Cognitive strategies outstand the Metacognitive cluster).

5. Discussion and Conclusion

To facilitate teachers' monitoring for ER activity, we explored a potential structure of a learners' model that can highlight their self-regulation strategies and extracted from multi-system interaction logs recorded in xAPI format. This shows an inventive use of interoperability standards to record learner's ER behaviors. 2 clusters of the students were identified and labeled as: Metacognitive & Cognitive Learners and Metacognitive Learners. We find the results can inform dashboard design and instructional support based on the visualized attributes of the cluster members, including the behavior of a single student and the behaviors of the clusters implying different learning strategies. This helps teachers monitor their behaviors, be aware of the implied strategies, and thus give learning strategy-based instruction.

We propose to design the visualization as 2 parts. On one hand, teachers can have a glance at students' actions by individual, which implies their behaviors in ER activity. On the other hand, teachers can also look at students' learning strategies by cluster, which implies their self-regulation strategies in ER activity. In terms of the former design, teachers can understand students' learning status, such as how many books were read, how many plans were made, how many reflections were shared, etc. As for the latter design, teachers can monitor students' strategies in a similar way. The visualizations may enable teachers to be aware that Metacognitive Learners only focused on organizing their reading plans but didn't read many books and share their reflection. Then, teachers can encourage these students to read more books so that cognitive strategies can be prompted.

The proposed design demonstrates how teaching and learning activities can be supported by xAPI when such experiences are distributed across various learning tools. The visualizations are expected to solve the difficulty of monitoring encountered in ER implementation. However, further evaluation is needed in future research. In addition, xAPI allows integrating student data from heterogeneous learning contexts with different software components. Hence, future research can also explore the possibility of extension to other learning activities.

Acknowledgements

This work is funded by JSPS KAKENHI 20K20131, 22H03902, NEDO JPNP18013, JPNP20006.

References

Carlson, R., Genin, K., Rau, M. A., & Scheines, R. (2013). Student Profiling from Tutoring System Log Data: When do Multiple Graphical Representations Matter?. In *Proceedings of the 6thInternational Conference on Educational Data Mining*.

Carrillo, R., Lavoué, É., & Prié, Y. (2016, April). Towards qualitative insights for visualizing student engagement in web-based learning environments. In *Proceedings of the 25th international conference companion on world wide web* (pp. 893-898).

Chang, A. C., & Renandya, W. A. (2017). Current practice of extensive reading in Asia: Teachers' perceptions. *The Reading Matrix: An International Online Journal*, *17*(1), 40-58.

Li, H., Majumdar, R., Chen, M. R. A., & Ogata, H. (2021). Goal-oriented active learning (GOAL) system to promote reading engagement, self-directed learning behavior, and motivation in extensive reading. *Computers & Education, 171*, 104239.

Lin, C. H., Hu, S. S., Lai, H. Y., Chiang, C. F., Tseng, H. C., & Cheng, Y. C. (2016, October). Visca: A dashboard system to visualize learning activities from e-learning platforms. In *International symposium on emerging technologies for education* (pp. 422-427). Springer, Cham.

Manso-Vázquez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2018). An xAPI application profile to monitor self-regulated learning strategies. *IEEE Access, 6*, 42467-42481.

Mustapa, A. M., Rahman, Z. A., Ghani, M. Z. A., Saad, M. F. M., & Mohamed, F. A. (2018). Qiraahbot's Prototype Development for An Extensive Reading Activity. *International Journal of Civil Engineering And Technology (IJCIET), 9*(9), 1494-1503.

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K., & Yamada, M. (2015, January). E-Book-based learning analytics in university education. In *International conference on computer in education (ICCE 2015)* (pp. 401-406).

Raga Jr, R. C., Raga, J. D., & Cariño, I. V. (2018). Visualizing Student Activity in Blended Learning Classroom by Mining Course Log Data. *International Journal of Learning and Teaching, 4*(1), 1-6.

Renandya, W. A., & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it. *Methodology in language teaching: An anthology of current practice*, 295-302.

Robb, T., & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language, 25*, 234-247.

Saint, J., Gašević, D., & Pardo, A. (2018, September). Detecting learning strategies through process mining. In *European Conference on Technology Enhanced Learning* (pp. 385-398). Springer, Cham.

Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces, 50*, 116-123.

Takii, K., Flanagan, B., & Ogata, H. (2021). An English Picture-book Recommender System for Extensive Reading Using Vocabulary Knowledge Map. In *Companion Proceedings of the 11th International Conference on Learning Analytics and Knowledge*.

Viberg, O., Khalil, M., & Baars, M. (2020, March). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 524-533).