Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

e-Book-based Learning Analytics in University Education

Hiroaki OGATA*, Chengjiu YIN, Misato OI, Fumiya OKUBO, Atsushi SHIMADA, Kentaro Kojima, Masanori Yamada
Faculty of Arts and Science, Kyushu University, Japan
*hiroaki.ogata@gmail.com

Abstract: This paper provides an overview of the Educational Big Data research project at Kyushu University, Japan. This project uses an e-book system called *BookLooper*. which allows students to browse e-books in Web browser, PC, mobile devices such as smartphone. This paper shows research issues in this project. Currently, about 2,700 first-year students are using the e-book system and approximately 2.2 million log data have been accumulated as of May 20, 2015. This paper describes why we introduce e-book in the University education and initial findings.

Keywords: Educational big data, learning analytics, e-book, learning log.

1. Introduction

Recently, digital textbooks or electronic textbooks (i.e., e-books) have been introduced to schools (Nakajima, Shinohara & Tamura, 2013) in many countries (e.g., Japan, Korea, and Singapore); this is especially true for K12 schools. For example, the Japanese Ministry of Education, Culture, Sports, Science and Technology compiled "The Vision for ICT in Education," a comprehensive policy that promotes utilizing ICT in education (MEXT, 2011). In this policy, the Japanese government also planned to introduce e-books in all K12 schools by 2020. In Korea, the research on e-books started in 1997 and in 2007, the e-book usage plan was announced by the Korean Education and Research Information Service (Shin, 2012).

Most of the pilot studies focused on introducing e-books at schools. However, very little attention has been paid to analyzing the e-book activity logs, although it is imperative to investigate how these logs can be used to improve e-book contents and the quality of learning and education. If the logs of K12 e-book learning activities are accumulated into the server, then *educational big data* can be accumulated. By analyzing the big data which includes information from e-books and LMS (learning management systems) and so on, learning activities can be supported and enhanced (Yin et al, 2014).

Many countries' e-book policies only focus on introducing the technology of e-books into K12 schools (Fang et al, 2012, Shepperd et al, 2008, Shin, 2012, Song, et al, 2007). This paper, however, tackles with introducing e-books at the university level. We believe there are several advantages that make it easier to introduce e-books in universities, as listed below:

(1) ICT skills: University students need to use ICT in their campus life, (e.g., to submit a report through LMS, to register courses in a web-based system, and to view the score of the courses in a web-based system). Thus, university students should have better ICT skills than K12 students.
(2) Internet accessibility: Some universities provide faster broadband internet access than in K12 schools. Thus, it may be easy for university students to download e-books on campus.
(3) Learning materials: Some professors have recently begun creating their own learning materials (e.g., by using PowerPoint and Keynote), revising them by themselves, and use them in their courses. Thus, it is easier to put these materials into the e-book system, as opposed to uploading entire books in K12 schools.
(4) Flexibility of course design and contents: In Japan, it is not easy for K12 teachers to change the course design and contents. However, professors in universities can change the design and contents

of their own course whenever they deem necessary. Thus, it is easier to integrate e-book activities in their courses; this is a very important factor for introducing e-books.

(5) Management of teaching and learning skills: Professors and university students should have the scientific skills needed to analyze their own teaching or learning log data. University students should also be more self-directed in their learning. With these points in mind, it is easier to introduce e-books and to utilize log data with university students rather than those in K12.

Due to the above reasons, we think that the university context provides an ideal test bed to introduce e-books. We would therefore argue that we need to test the effectiveness of e-books and the utilization of e-book logs at the university level first.

Furthermore, regarding the ICT environment at Kyushu University, the University has implemented a BYOD (Bring Your Own Devices) policy since 2014, and provides a broadband wireless Internet access (300 Mbps) in all lecture rooms on campus. Every weekday, they bring their own laptops to attend their courses. In this paper, we describe the overview of the system and initial findings.

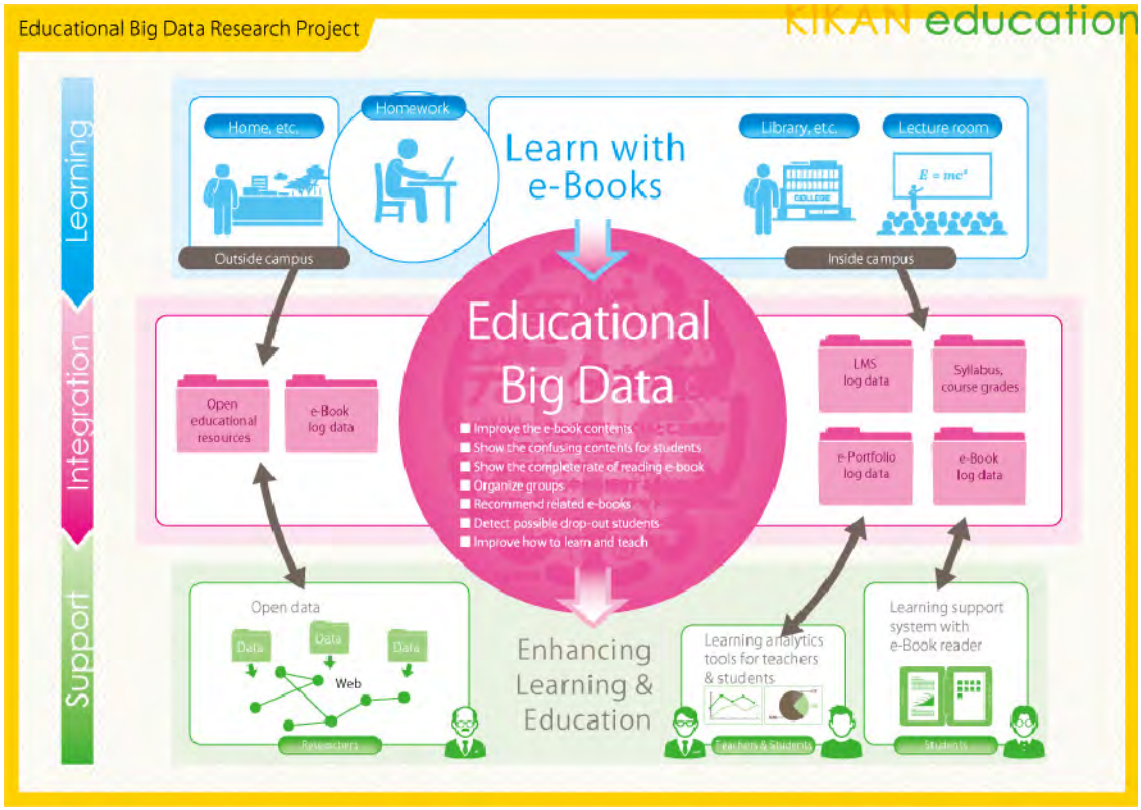


Figure 1: Overview of the research project.

2. M2B system

We define educational big data as the integration of several kinds of data that includes the educational and learning activities of teachers and students inside and outside school. We are using Moodle as our LMS, Mahara as the e-Portfolio system, and BookLooper as the e-Book system. We call the three systems the "M2B (Mitsuba) system." All of the logs from the M2Bsystem are integrated and analyzed. For example,

(1) The teachers use Moodle to manage student attendance, provide quizzes, and receive reposts. These data and the access logs on Moodle are also used as a part of the educational big data.

(2) Both the teachers and students wrote e-portfolio after each lecture by using Mahara. E-portfolio is useful in evaluating a course qualitatively and to improve the context of the e-book.

(3) BookLooper[1] is a commercial product designed by Kyocera Maruzen Systems Integration Co., Ltd. This system provides a cloud service; the e-books are then managed in the cloud based on digital right management (DRM). Users can download these materials using the BookLooper viewer. BookLooper provides a client viewer application that works on several computing devices including Windows, Macintosh, iPhone/iPad, and Android devices.

Functions of e-books

The e-books are managed in the bookshelf. If users select a book in the bookstore, the book will be downloaded into the bookshelf. The users then choose the book in the bookshelf in order to read it in the viewer. If the book is downloaded on the users' computing device, users can view it without an internet connection. In the viewer, the user can use functions such as next page, previous page, bookmark, underline, annotation, and keyword search. All the activities in the viewer are recorded on the device with a time stamp, and the logs are sent to the server when an internet connection becomes available. The e-books of BookLooper can be also opened by clicking on a link in Moodle. Then, the system can know when and for which course the e-book was used. This information is very useful and important in analyzing students' learning activities.

(A) Bookstore (B) Bookshelf (C) Viewer

Figure 2: e-Book interface.

Educational materials for e-books

Most professors are using paper-based books for their courses, which are distributed by various publishing companies. Thus, it is difficult to make these books available on the e-book system due to copyright issues. We are thus currently using the lecture slides (PowerPoint or PDF) that are created by the teachers. BookLooper converts the lecture slides into an original e-book file; the students can then download and view all e-book files once the materials are stored on the BookLooper server. Since BookLooper is based on a special file format and not the ePub format, there are some limitations. For example, BookLooper cannot store animation, movie, and webpage URLs.

e-Book logs

Logs for this analysis were collected from 98 first-year students via BookLooper. These students took an information science course in the second semester of the 2014/2015 school year at Kyushu University (table 1). Figure 3 shows sample e-books logs. Note that the data from 15 students who did not take the midterm or the term-end examination of the course were discarded from the further analyses. There are many types of operations in logs, for example, OPEN means that the student opened the e-book file and NEXT means that the student clicked the next button to move to the subsequent page. PORTRAIT signifies that the student turned the computing device into the portrait position.

[1] More information (in Japanese) can be found at http://www.kccs.co.jp/ict/cloud-booklooper/

Devices	User ID	e-book ID	Title (e-book)	Operation	Page	Status	Marker Start	Marker End	Date	Time	Duration
AIR	xxxxxxxxxx	00000000NKFS	A-10-通信路符	OPEN	50	0	0	0	2014-11-12	08:54:35	3
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	OPEN	0	0	0	0	2014-11-12	08:54:42	0
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	PORTRAIT	1	1	0	0	2014-11-12	08:54:46	2
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	NEXT	2	1	0	0	2014-11-12	08:54:48	1
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	NEXT	3	1	0	0	2014-11-12	08:54:49	6
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	CLOSE	3	0	0	0	2014-11-12	08:54:55	68
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	OPEN	3	0	0	0	2014-11-12	08:56:03	5
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	NEXT	4	1	0	0	2014-11-12	08:56:08	1
AIR	xxxxxxxxxx	00000000NKG4	A-11-暗号	NEXT	5	1	0	0	2014-11-12	08:56:09	8

Figure 3. Samples of e-book logs.

Semester B starts from April and finishes at the end of July 2015 as shown in table 1. Thus, approximately 6,600,000 log data will be collected. If the data is integrated with Moodle and Mahara data, we may accumulate educational big data.

Table 1: Collected data by using e-book as of May 20, 2015.

	Semester A	Semester B (will end on July 31)
Term period	Oct 1, 2014 – Feb. 6, 2015	April 13, 2015 – May 20, 2015
# of log data	581,352	2,231,375
# of students	306	2,687
# of courses	5	38
# of e-books	148	183

3. Research agenda

There are many areas of focus in this research project, which are described as follows:

(1) Improve learning materials: We can ask students to underline or mark unknown words in the e-books. According these logs, the system can detect words students cannot fully understand. The system can then show the information to the teacher and the teacher will add more explanation about those words in the e-book.

(2) Analyze learning patterns: According to the e-book logs, the system may detect a student's learning pattern (e.g., learning style/cognitive style). By doing so, the system can promote students' learning and increase their success in class.

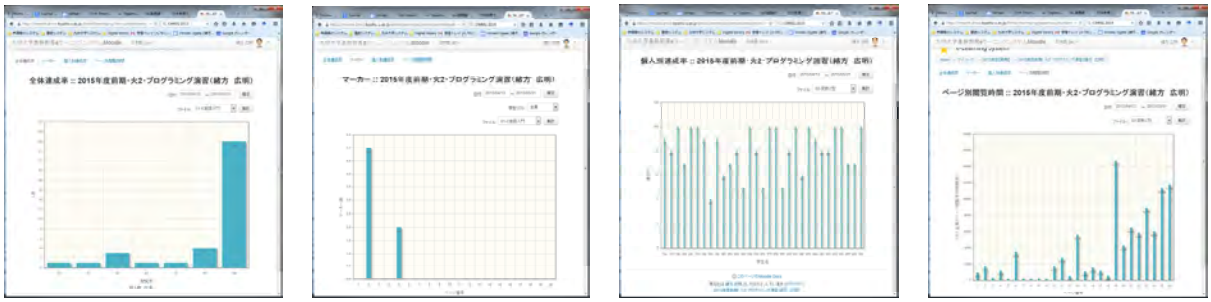
(3) Detect students' comprehensive level: If the system can integrate e-book logs with the results of quizzes or examinations in Moodle, then the system can show teachers the status of students' comprehension level. If students do not read the e-book or cannot answer a quiz on the e-book's content, they may not understand the context of e-book.

(4) Predict final grades: By using the M2B logs, the system may predict a student's final course grade and identify any students at risk of dropping out/failing. The teacher can then issue warnings to students early in the course.

(5) Recommend e-books: By analyzing the M2B system logs, the system can detect a student's interest and recommend e-books to promote active learning. The students' interests can also be used to organize students into groups where the members have different interests.

4. Current status

We have developed a Moodle plugin program to show the results of the e-book log analysis. The graph is slightly different based on whether the user is a teacher or a student. More specifically, teachers can see all of the students' logs individually. In contrast, students can only see their own name and logs in the graph.



(A) Completion rate. (B) Markers per page. (C) Completion rate per student. (D) Duration per page.
Figure 4: Analysis results of e-book logs shown in Moodle.

Figure 4 shows the graphs in Moodle (teacher's view):

(A) Completion rate: X-axis is the rate (from 0 to 100%) and y-axis is the number of students. From this graph, the teacher can understand how many students have already read e-book before each class. The completion rate (*CR*) is calculated by:

$$CR(\%) = \frac{\#\text{ of the pages in the e-book that the student read}}{\#\text{ of the pages in the e-book used for a given course}} \times 100$$

A certain page is considered to be "read" if a student views the page for more than the given time period, e.g., 10 seconds, which the system administrator can define. If most of the students read the e-books, then the teacher can reduce the time spent on explaining the pages and dedicate more time for group work, problem-based learning, or other exercises. Otherwise, the teacher should refer to graph (D) and explain any pages students did not read in more detail.

(B) Markers per page: X-axis is the page of the e-book and y-axis is the number of marks students made in the page. The teacher told the students to mark any unknown words in the e-book. By this graph, the teacher can thus understand which words were difficult for students. Therefore, during the course, the teacher can elaborate more on any topics the students had marked. Also, the teacher can improve the learning material by adding more explanations in it.

(C) Completion rate (CR) per students: X-axis shows the student ID and y-axis is the CR of the student. This graph can only be viewed by teachers.

(D) Duration per page: X-axis is the page in the e-book file and y-axis is the time (in seconds) that students viewed the page for. If students spend a lot of time on a certain page, the teacher should explain this page more and add further explanations on the page to improve the e-books.

Initial findings using the data are as follows:

We observed two rules from the data in semester A.

(1) If the student viewed the e-book for less than 5 minutes in the first four lectures of a course, then the score of the student will be less than 79 out of 100 in 93.8% probability.

(2) If the student viewed the e-book for more than 5 minutes in the first four lectures of a course, then the score of the student will be more than 80 out of 100 in 100.0% probability.

5. Conclusions

This paper describes a research project that accumulated and analyzed educational big data by using a M2B system (i.e., Moodle, Mahara, BookLooper). From the initial experiment, this system may predict the final score if the course in the first four lectures by using e-book logs. In future work, we will allow teachers and students to download their own data; the system will provide them with data analysis tools to manage their learning and teaching skills. From a technological point of view, we will tackle research issues such as data integration, real time data mining, visualization, recommendation, and predictions. In addition, we will integrate e-book and SCROLL (Ogata, et al, 2011) in order to enhance learning experiences.

Acknowledgements

The research have been supported by “Research and Development on Fundamental and Utilization Technologies for Social Big Data” (178A03), the Commissioned Research of the National Institute of Information and Communications Technology (NICT), Japan.

References

Fang, H., Liu, P., Huang, R. (2011). The Research on E-book-oriented Mobile Learning System Environment Application and Its Tendency, International Conf. on Computer Science and Education, pp.1333-1338.

MEXT, Japanese Ministry of Education, Culture, Sports, Science and Technology (2011). “The Vision for ICT in Education”, http://www.mext.go.jp/b_menu/houdou/23/04/__icsFiles/afieldfile/2012/08/03/1305484_14_1.pdf

Kim, J. H.-Y., & Jung, H.-Y. (2010). South Korean digital textbook project. *Computers in the Schools, 27*, 247-265.

Nakajima, T., Shinohara, S., & Tamura, Y. (2013). Typical functions of e-Textbook, implementation, and compatibility verification with use of ePub3 materials. *Procedia Computer Science*, *22*, 1344–1353.

Ogata, H., Li, M, .Bin, H., Uosaki, N., El-Bishoutly, M. & Yano, Y. (2011). SCROLL: Supporting to share and reuse ubiquitous learning logs in the context of language learning. *Research and Practice on Technology Enhanced Learning*, 6 (3), 69-82.

Shimada, A., Okubo, F., Yin, C-J., and Ogata, H. (2015). Informal Learning Behavior Analysis Using Action Logs and Slide Features in E-textbooks. *Proceedings of the 15th IEEE International Conference on Advanced Learning Technologies (ICALT 2015)*, in print

Shepperd, J. A., Grace, J. L., & Koch, E. J. (2008). Evaluating the electronic textbook: Is it time to dispense with the paper text? *Teaching of Psychology*, *35*(1), 2–5.

Shin, J. H. (2012). Analysis on the digital textbook’s different effectiveness by characteristics of learner.*International Journal of Education and Learning*, *1*(2), 23–38.

Song, H. D., Jun, J. S., & Ryu, J. H. (2007). *The effects of digital textbooks in student learning*. Seoul, South Korea: Seoul Metropolitan Board of Education.

Yin, C., Okubo, F., Shimada, A., Kojima, K., Yamada, M., Fujimura, N. and Ogata, H. (2014). Smart phone based data collecting system for analyzing learning behaviors, *Proc of International Conference of Computers in Education 2014, 575–577.*

Yamada, M., Yin, C., Shimada, A., Kojima, K., Okubo, F. and Ogata, H. (2015). Preliminary Research on Self-regulated learning and learning logs in a ubiquitous learning environment, *Proceedings of the 15th IEEE International Conference on Advanced Learning Technologies*, 93-95.