IIAI Open Conference Publication Series
IIAI Letters on Institutional Research
Vol. 003, LIR116
DOI: https://doi.org/10.52731/lir.v003.116

Establishment of an IR System for Institutional Decision-Making in Kyushu Institute of Technology

Tetsuya Oishi *, Hirofumi Fukumaru *,
Akihiro Hayashi *, Hiroshi Sakamoto *

Abstract

The Kyushu Institute of Technology (Kyutech) has established a system that enables students and teachers to reflect on their learning history and complete self-evaluations using e-portfolio systems. However, institutional research (IR) for institutional decision-making has not yet been introduced, and an Educational IR Support Group was established in a learning-teaching center in FY2022 to strengthen IR for education. In addition to the e- portfolio systems, data have been accumulated in various other systems; however, these data are managed by different departments, and there is currently no system in place to utilize the data comprehensively. We introduced an extract transform load (ETL) tool to combine and analyze data and then began building an IR system for education. In this study, we first provide an overview of IR and explain how it should be conducted. Then, we introduce the IR system for education being promoted and finally describe future work.

Keywords: Decision Making, Educational Data, Information Support Cycle, Institutional Research for Education, Visualization

1 Introduction

The Kyushu Institute of Technology (Kyutech) began paper-based self-assessment for students in 2003, and after the School of Computer Science and Systems Engineering received JABEE [1] certification in 2005, self-assessment was expanded to the entire university. In 2007, the development of a learning self-assessment system, an e-portfolio system for students, began. In 2009, the e-portfolio system, which can also record extracurricular activities, was put into university-wide operation. The system has been established throughout the university and is continually improved [2]. While there are systems for collecting and feeding educational data back to individuals at our university, there are no systems that sup- port institutional decision-making. To strengthen institutional research (IR), an Educational IR Support Group was established in the learning-teaching center in FY2022. This paper provides an overview of IR, describes the IR at Kyutech, and discusses future research.

2 Institutional Research

Due, in part, to the recent decline in the population of 18-year-olds, many higher education institutions have begun to focus on IR to sustain their respective institutions. Each institution now has

* Kyusyu Institute of Technology, Fukuoka, Japan

its own IR organization and functions. Although there is no fixed method for conducting IR, a concept called the information support cycle [3] is employed as a model. The information support cycle comprises five phases that are repeated to support institutional decision-making.

1. Identification of issues and needs
2. Data collection and accumulation
3. Data reconstruction and analysis
4. Data reporting
5. Decision making

The cycle means that some decisions not only identify new issues and needs, but also lead to new decisions. “2. Data collection and accumulation”, “3. Data reconstruction and analysis”, and “4. Data reporting” are the tasks of IR practitioners, while “1. Identification of issues and needs” and “5. Decision making” are the tasks of the university executives running each institution. IR is a mechanism that supports stakeholders’ decision-making, especially when dealing with educational data, and is sometimes called educational IR. This shows that the primary task of IR practitioners is to work with the data. However, this often does not include data management because it depends on the operations that generate the data. These operations have their own specific methods for handling data. IR practitioners are not involved in data management, but only in providing data. The following section describes the tasks performed by IR practitioners.

2.1 Data Collection and Accumulation

By classifying the data in higher education institutions, unstructured data, such as text, and structured data are expressed as numerical values [4]. The structured data managed by each department in charge are collected and stored in a system called a **data lake** by directly referencing files such as CSV or databases.

2.2 Data Reconstruction and Analysis

Data collected in the data lake is combined using an extract transform load (**ETL**) tool and restructured to facilitate analysis for data reporting. At this time, data that are easier to analyze are stored in a system called a **data warehouse**. It was found that the longer the IR experience, the more important the ETL tool used for data reconstruction, especially for the ETL tool used for data reconstruction, although inexperienced IR practitioners tend to be less interested in the data reconstruction part of the process [5].

2.3 Data Reporting

IR practitioners visualize the data stored in the data warehouse and report them to decision- makers using graphs, etc., as necessary. The data to be stored in the data warehouse must include various elements for visualization from multiple angles. A **business intelligence** (BI) tool may be introduced to increase the real-time nature of reporting and streamline data visualization.

3 Educational IR in Kyutech

Kyutech has a Learning and Teaching Center to support its educational work. We describe this center in Section 3.1, the educational support systems in Section 3.2, and an educational IR Support Group in Section 3.3.

3.1 Learning and Teaching Center

The Learning and Teaching Center consists of four groups: the Educational IR Support Group, the Educational DX (Digital Transformation) Support Group, the FD (Faculty Development) Support Group, and the Technical Support Group. These four groups work together to manage the center. In particular, we describe the educational support systems operated by the Learning Support Group, which was the predecessor of the Educational IR Support Group.

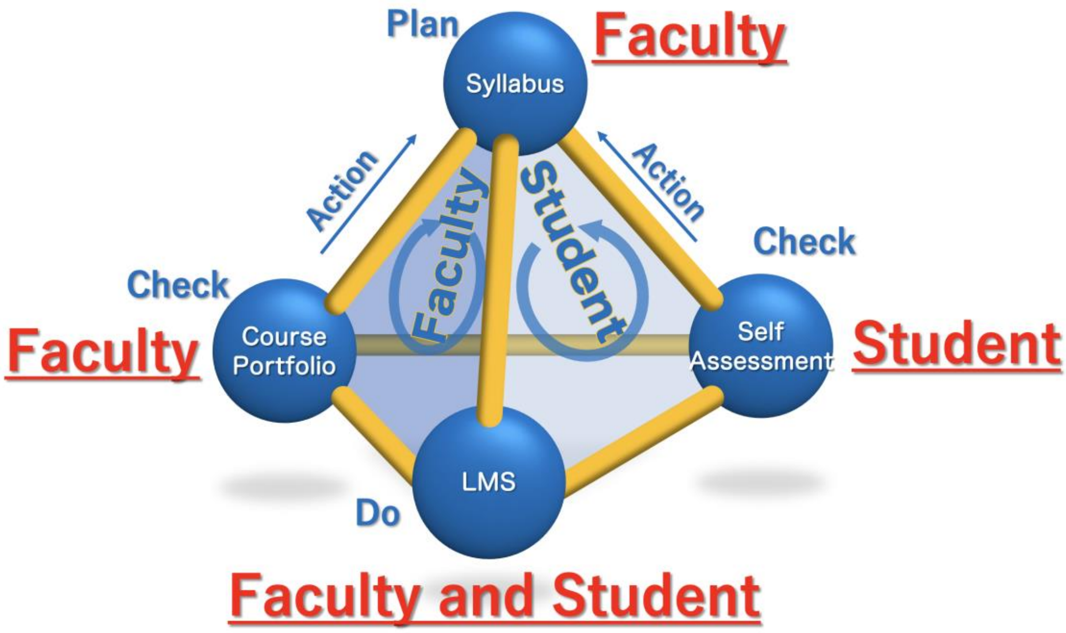

Figure 1: Correlation Chart of Education Support Systems

3.2 Education Support Systems

Kyutech has four educational support systems, as shown in Figure 1: a **Syllabus System**, a Learning Management System (**LMS**), and two e-portfolio systems (a **course portfolio** for faculty and a **self-assessment** system for students). We describe each of them below.

3.2.1 Syllabus System

The syllabus system, located upside in Figure 1, allows faculty to show not only class outlines and lesson plans, but also the positioning of classes based on the Diploma Policy and the Curriculum Policy. In addition, the syllabus system can also set class achievement objectives and evaluation criteria. The syllabus system corresponds to “Plan” of the PDCA, an acronym for “plan, do, check, action.”

3.2.2 LMS

The LMS, located at the bottom of Figure 1, will share educational information linked to the syllabus, accumulate learning outcomes that support the implementation of the educational plan, and provide students with learning materials. Our university has adopted Moodle 4.1 [6] as its

LMS. LMS corresponds to “Do” of the PDCA.

3.2.3 E-Portfolio Systems

Kyutech has two e-portfolio systems: the Course Portfolio, located on the left side of Figure 1, and the Self-Assessment system, located on the right side of Figure 1. The Course Portfolio is an e-portfolio for faculty, which enables them to analyze grades, analyze student self-assessments, review achievement goals, check student learning outcomes, and evaluate educational effectiveness, corresponding to “Check” of the PDCA for faculty. On the other hand, the Self-Assessment system is an e-portfolio that allows students to set learning goals, self-evaluate their achievement, check their grade history, reflect on their learning, and record both regular and extracurricular activities, which corresponds to “Check” of the PDCA for students.
Reflecting the contents checked by the faculty and students in their respective portfolio systems in the syllabus system is an activity corresponding to “Action” in the PDCA.

3.3 Educational IR Support Group

As explained in the previous section, Kyutech has gathered data on education and has a mechanism in place to support individual decision-making not only by faculty but also by students. However, there is no mechanism to support institutional decision-making. Therefore, the Educational IR Support Group was established to support institutional decision- making in 2022.
At Kyutech, we started by introducing the ETL tool described in section 2.2 to combine the educational data within our university into an easily analyzable form. Ideally, there should be a mechanism to collect and store data using a data lake, as described in section 2.1, and a mechanism to store reconstructed data using a data warehouse, as described in Section 2.2; however, in our university, the functions of the IR system were realized by including these functions in the server running the ETL tool. We introduced a semiautomatic visualization mechanism using the statistical analysis software R [7] to visualize the combined data using the ETL tool.
The current server configuration that realizes the semiautomatic visualization mechanism is shown in Figure 2. (1) The data providers transfer the data from servers that cannot be referenced directly to the shared server, and (2) IR practitioners manually store these data in the storage equipped in ETL server. Not only (3-1) data in the storage equipped in the ETL server but also (3-2) data from servers that can be referenced directly are combined using the ETL tool. These data are (4-1) analyzed and visualized using the statistical analysis software R and (4-2) converted to PDF using Latex. (4-3) The data combined with the ETL tool, (4-4) the data visualized by R, and (4-5) the data converted to a PDF by Latex are stored in the storage equipped in the ETL server. Then (5) IR practitioners store them on a shared server. (6) Data providers and decision-makers can retrieve reported data from a shared server.
At the time of writing, we contributed to institutional decision-making by realizing a semiautomatic visualization of grade distribution and withdrawal status from the perspective of IR for education.

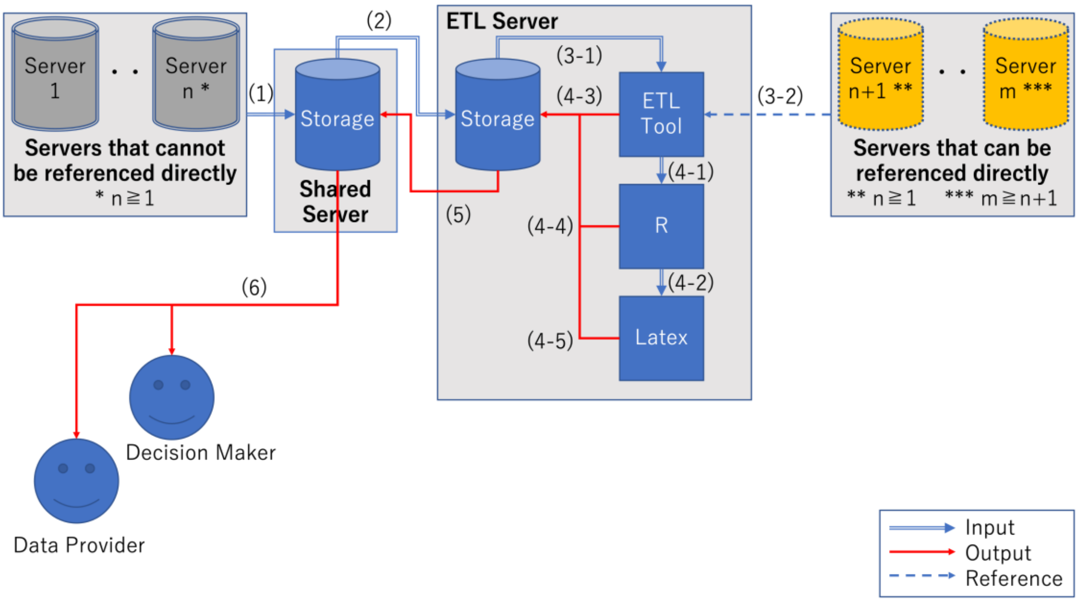

Figure 2: Current server configurations

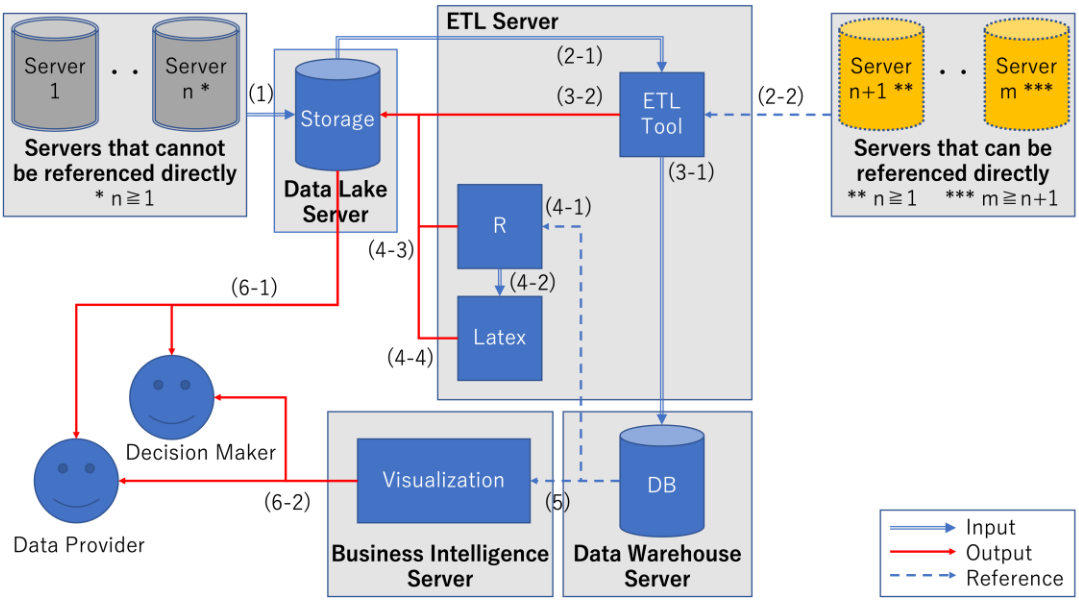

Figure 3: Future server configurations

4 Future Work and Expected Effects

In this section, we discuss future work and expected effects.

4.1 Future Work

A server configuration that can realize an automatic visualization mechanism is shown in Figure 3.

(1) The data providers transfer data from servers that cannot be directly referenced to the storage equipment in the data lake server. This paragraph is equivalent to “2. Data collection and accumulation” in section 2.
Not only (2-1) data in the storage equipped in the data lake server but also (2-2) data from servers that can be referenced directly are combined using the ETL tool. At this time, because the storage equipment in the data lake server will be mounted on the ETL server, we will be able to use the data as if they were in the storage equipment in the ETL server. Moreover, the data combined by the ETL tool are not only (3-1) stored in the database equipped in the data warehouse server, but also (3-2) again stored in the storage equipped in the data lake server. (4-1) The data from the database equipped in the data warehouse server were analyzed and visualized using the statistical analysis software R, and (4-2) converted to PDF using Latex. (4-3) The data visualized by R and (4-4) the data converted to PDF by Latex were again stored in the storage equipped in the data lake server. This paragraph is equivalent to “3. Data reconstruction and analysis” in section 2. The reason why R and Latex exist in the ETL server in spite of the Business Intelligence function described below is to enable flexible visualization for urgent requests.
On the other hand, (5) the data from the database equipped in the data warehouse server can be interactively visualized on the business intelligence server. Data providers and decision-makers can not only (6-1) retrieve reported data from the data lake server but also (6-2) retrieve visualized data on demand using the visualizing system working in the business intelligence server. This paragraph is equivalent to “4. Data reporting” in section 2.

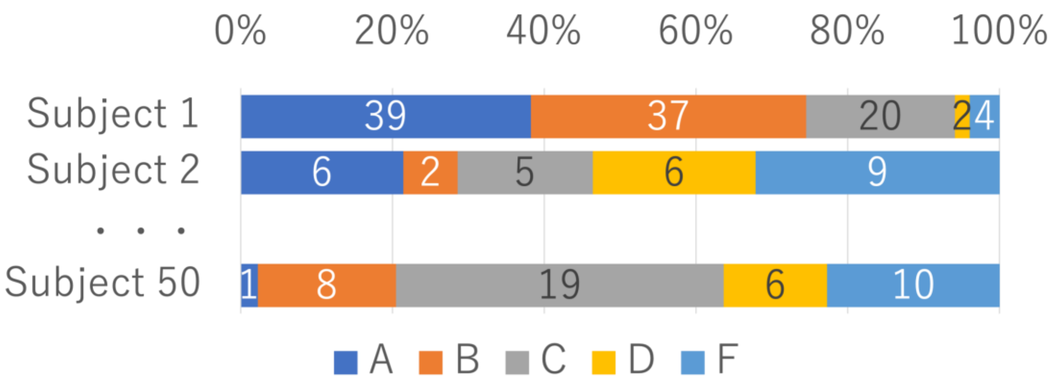

Figure 4: Example of Grade Distribution Report

4.2 Expected Effects

At Kyutech, we provided decision-makers with the grade distribution shown in Figure 4. This section uses this visualization as an example to illustrate the expected effects.
Kyutech has approximately 2,500 subjects. For simplicity, we assume that 2,400 sub- jects exist. At Kyutech, we present the grade distribution on four separate occasions. This means that the grade distribution of 600 subjects was presented each.
The data provider reported that it took 2 hours to extract the original data for grade distribution. Given that it took approximately 10 minutes to visualize 50 subjects per page as shown in Figure 4, it would have taken 2 hours to visualize 600 subjects, or 12 pages. You can see that these tasks will be performed four times a year, so a total of 16 hours will be required as shown Table 1 (1).

Table 1: Time Required to Report a Grade Distribute

	Times Required One Session	Four Sessions	Time Reduced Compared to (1)
(1) Past	4 hours	16 hours	-
(2) Now	2.25 hours	9 hours	7 hours
(3) Near Future	0.25 hours	1 hour	15 hours
(4) Future	0 hours	0 hours	16hours

Next, we describe the visualization using the current server configuration shown in Figure 2. The data extraction time of the data provider was maintained at 2 hours, but the visualization task was reduced to 0.25 hours because automation by ETL tool and R was realized. At 2.25 hours of work per session, this translates to 9 hours of work per year, which means a reduction of 7 hours of work per year, as shown in Table 1 (2).
By making the database available for direct reference, the time required by the data provider to extract data is eliminated. Then, the time for the data provider to extract data becomes 0 hours, and only 0.25 hours of visualization work is required per operation. In other words, 1 hour was required per year for this work, as shown in Table 1 (3).
Moreover, we describe the visualization using a future server configuration, as shown in Figure 3. The functions discussed thus far only require the cost of maintaining an ETL server. To realize these functions after this point, additional maintenance costs for data warehouses and BI servers will be required. Decision makers can automatically see the visualized data; therefore, their work time becomes zero, as shown in Table 1 (4).
We explained that one visualization task saved 16 hours. In other words, if there were n visualization tasks, it would be possible to save $16n$ hours.

5 Summary

This study provides an overview of IR and its status at our university. Our IR system is still in its developmental stage, and we are constructing an IR system through a trial- and-error process. In the future, we will start the construction of a data lake, and while referring to examples of IR at other universities, we will consider the introduction of a data warehouse and business intelligence tools based on our university's situation. Currently, we are attempting to use this system to analyze educational data, but in the future, we would like to combine admissions data and campus life data for further analysis. In addition, as the number of data types and stakeholders increases, management is expected to become more difficult. Therefore, care must be taken to avoid complicated management.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 21K02653.

References

[1] “JABEE,” https://jabee.org/en/ (accessed 31st March 2023).

[2] Hiroshi Sakamoto, Hirofumi Fukumaru, Takashi Miyaura, Kazunori Nishino, Akihiro Hayashi, “Use Case of Visualization of Learning Achievement Using e-Portfolio - Construction of Internal Quality Assurance of Education and Formation of Consortium by Diverse Educational Institutions -,” Joint Forums between University e-Learning Association and The Japan Association for Development Education 2019, 2020, pp. 31-32.

[3] R.D.Howard (ed.), National Institution for Academic Degrees and Quality Enhancement of Higher Education, “Institutional Research: Decision Support in Higher Education,” Tamagawa University Press, 2012.

[4] Yoshiharu Aioi, Shuntaro Iseri, Shotaro Imai, Tetsuya Oishi, Saori Okada, Nobuhiko Kondo, Toru Sugihara, Shintaro Tajiri, Mio Tsubakimoto, Keita Nishiyama, Takeshi Matsuda, Masao Mori, “Data Collection,” Standard Guidebook for University IR， Japan Association for Institutional Research, 2022, pp.122-126.

[5] Tetsuya Oishi, “What is the Essential Curriculum for IR in Japan?,” Proceedings - 2022 11th International Congress on Advanced Applied Informatics, IIAI-AAI 2022, LIR009, 2022, pp. 1-5.

[6] “Moodle,” https://moodle.org/ (accessed 4th April 2023).

[7] “RStudio,” https://posit.co/download/rstudio-desktop/ (accessed 31th March, 2023).