Learning Analysis infrastructure to support optimal learning based on IRT and Video Enrollment Data

Masahiro SAKAMOTO, Hiroshima City University Graduate School of Information Sciences and Technology, Japan
Yukihiro MATSUBARA, Hiroshima City University Graduate School of Information Sciences and Technology, Japan
Kousuke MOURI, Hiroshima City University Graduate School of Information Sciences and Technology, Japan
di65006@e.hiroshima-cu.ac.jp

Abstract: Currently, local governments in Japan need DX (=Digital Transformation) promotion personnel. However, even today, Japan's civil service system still employs mainly legal and human personnel, resulting in a need for more necessary human resources. Therefore, the demand for DX human resource development for local government officials is rapidly increasing. In response to this, we have been engaged in human resource development using e-learning materials based on Moodle for three years since 2020, targeting approximately 9,500 employees of the Hiroshima City Office, where I work, and through formative evaluation of the materials, we have succeeded in creating materials at a level where all employees who took the course expressed a desire to promote DX. The company succeeded in developing educational materials so that all staff members who took the course were willing to promote DX. However, only about 5% of all staff members completed the course, reaching the limit of the improvement method in formative evaluation from those who completed the course. During the three-year effort, we also identified a wide gap in the IT skills of local officials and that we cannot provide optimal educational materials for 95% of the staff. In this study, we created multiple-choice questions corresponding to the DX literacy standard set by the Ministry of Economy, Trade and Industry (METI) in December 2022, which we regarded as a complete mastery map, and optimized them using IRT. The objective was to develop a model for measuring the DX literacy of local government officials and an e-learning course that provides individualized and optimal learning based on the measurement results. Through these efforts, we would like to establish a method of correlating skill surveys using IRT (=Item Response Theory) with e-learning course attendance data to overcome the limitations of the technique of creating e-learning courses and improving them through formative evaluation and to present a method to overcome the limitations of formative assessment in the ID (=Instructional Design) field. We hope to offer a strategy to overcome the limitations of formative evaluation in ID.

Keywords: DX, e-learning, IRT, Local government, DX literacy standard, Video Analysis, Skill Survey

1. Introduction

1.1 LMS Infrastructure

An e-learning course for Hiroshima City employees was created in the environment shown in Figure 1, with construction starting in October 2019 and course offerings beginning in March 2020. Furthermore, based on the load data from the two courses initially provided, SSL (=Secure Socket Layer) support and Vimeo (= Vimeo is an online platform for sharing and watching videos. There most significant difference between Vimeo and YouTube is that

creators pay a usage fee. There is no advertising.) plan changes were made. As a result, five classes were delivered to Hiroshima City employees for three years until March 2023, functioning without a single failure as an e-learning infrastructure for 20,000 viewers, 3,000 students, and 800 completed courses; this study allowed us to verify the availability of the LMS information infrastructure construction in terms of required CPU resources, for example.
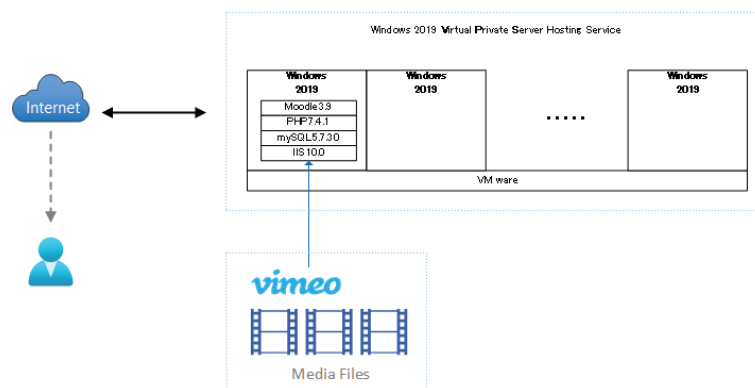

Figure 1. e-Learning Infrastructure.

Figure 2. e-Learning Courses.

1.2 Results of e-learning Course offerings

Based on about 500 survey responses from participants who completed the course and formative feedback from interviews with about 20 participants, we spent three years improving the course materials to create a system that is "easy for participants to understand." (Table 1).

Table 1. Responses obtained in formative evaluation

	FY2021	FY2022
Fairly inaccessible	8%	2%
A little difficult to approach	25%	6%
It was no different from any other material.	21%	8%
It was relatively easy to approach.	30%	44%
I found it very approachable.	16%	40%

As an example of improvement based on formative evaluation, we worked on improvements such as carefully explaining IT terms in the videos.

In FY2020, as shown in Table 2, videos longer than 14 minutes tended to increase the number of students who left the course. In addition, in the formative evaluation by participants, some said that one class was too long for them to study during their workday, which interfered with their work.

Table 2. Video viewing time (FY2020)

Year	Course Name	Impression	View	Completion	Cmpletion ratio	Video time	Average	Unwatched
FY2020	Information security training (Users Section) [Lecture 1]	4025	404	238	58.91%	0:22:31	0:16:59	0:05:32
	Information security training (Users Section) [Lecture 2]	3998	291	177	60.82%	0:22:58	0:18:17	0:04:41
	Information security training (Users Section) [Lecture 3]	3995	227	160	70.48%	0:14:11	0:12:24	0:01:47
	Information security training (Users Section) [Lecture 4]	3997	248	173	69.76%	0:01:51	0:01:34	0:00:17
	Information security training (Users Section) [Lecture 5]	3998	229	124	54.15%	0:27:53	0:22:46	0:05:07
	Information security training (Managament section) [Lecture 1]	1309	183	110	60.11%	0:23:05	0:17:14	0:05:51
	Information security training (Managament section) [Lecture 2]	1307	144	69	47.92%	0:41:37	0:27:06	0:14:31
	Information security training (Managament section) [Lecture 3]	1282	97	64	65.98%	0:18:23	0:15:03	0:03:20
	Information security training (Managament section) [Lecture 4]	1280	88	49	55.68%	0:39:53	0:29:20	0:10:33
	Information security training (Managament section) [Lecture 5]	1276	81	50	61.73%	0:28:37	0:23:01	0:05:36

Using these results and the formative evaluation as a reference, the course was redesigned in FY2021 so that all videos would be approximately 10 minutes long. As a result, as shown in Table 3, unwatched time was significantly reduced, and the number of students who were likely to leave the course due to the length of the videos was reduced.

Table 3. Video viewing time (FY2021)

Year	Course Name	Impression	View	Completion	Cmpletion ratio	Video time	Average	Unwatched
FY2021	2022 teaser	6416	225	166	73.78%	0:00:17	0:00:13	0:00:04
	DX starting from Information Security (General) [Lecture 1]	775	417	280	67.15%	0:10:32	0:08:40	0:01:52
	DX starting from Information Security (General) [Lecture 2]	530	283	209	73.85%	0:11:27	0:10:14	0:01:13
	DX starting from Information Security (General) [Lecture 3]	475	268	194	72.39%	0:12:51	0:11:45	0:01:06
	DX starting from Information Security (General) [Lecture 4]	472	251	191	76.10%	0:10:32	0:09:29	0:01:03
	DX starting from Information Security (General) [Lecture 5]	435	250	157	62.80%	0:11:17	0:09:39	0:01:38
	DX starting from Information Security (Management) [Lecture 1]	307	189	143	75.66%	0:10:16	0:08:51	0:00:20
	DX starting from Information Security (Management) [Lecture 2]	260	165	111	67.27%	0:09:38	0:08:17	0:01:25
	DX starting from Information Security (Management) [Lecture 3]	215	129	95	73.64%	0:08:33	0:07:37	0:01:21
	DX starting from Information Security (Management) [Lecture 4]	185	120	85	70.83%	0:13:01	0:11:10	0:00:56
	DX starting from Information Security (Management) [Lecture 5]	195	119	89	74.79%	0:09:25	0:08:18	0:01:51
	DX starting from Information Security (Management) [Lecture 6]	205	117	53	45.30%	0:13:03	0:11:36	0:01:07

By the way, the Information Policy Department of the City of Hiroshima provides consultation on DX promotion. As a result of the provision of e-learning courses and improvement of course content, the number of staff consulting on DX promotion has repeatedly doubled since the start of training compared to the previous year in an evaluation using Kirkpatrick's four-level evaluation method. The number of staff starting DX promotion, in other words, reaching Level 3, has also increased (Table 4).

Table 4. Number of Consultations

Year	Number	Comment
2018	28	
2019	30	
2020	23	
2021	49	Results for 2020
2022	104	Results for 2021

However, in this environment, it is not possible to use methods other than formative evaluation, although it is possible to grasp the students' quiz-taking status in each course, which can be obtained through Moodle, the results of questionnaires, and the viewing time of each lecture video provided on Vimeo.

This made it impossible to obtain data to address students who did not cooperate with the formative evaluation or who dropped out of the program, making further improvement unlikely (Table 1).

1.3 Organization of conventional methods and Issues

The program's primary goal is to design an efficient learning experience for the participants based on the data by correlating test results with their attitudes toward the course.

The Hiroshima City employees targeted for the course were widely distributed between the ages of 18 and 65, including some who had graduated from information

technology universities and some who had worked in the private sector, and there were significant differences in their abilities and IT experience.

In the formative evaluation of a previous study, we analyzed the results from the perspective that "older employees are less comfortable with IT." Still, we were able to determine that there was a relationship between pre-training test scores and ease of use rather than age.

However, since the pre-test was merely a questionnaire about the language information domain set based on the course content, the validity of the test itself and the accompanying course evaluation could not be statistically processed.

Several private operators currently offer tools for conducting skills surveys that comply with the DX Literacy Standards. Still, they are "fixed question sets" and require about one hour to take the test. Furthermore, when we inquired with the providers about the content, they said they did not evaluate the validity of the questions. In this regard, by using IRT, we will create a test corresponding to the learning items listed in the DX Literacy Standards and evaluate its validity.

Furthermore, instead of having the respondents answer the entire massive database of questions, the test should be an adaptive test to measure their understanding of the DX Literacy Standards in a 15-minute test that can be administered between work hours. Finally, through these efforts, we hope to quantitatively measure the ability of "DX human resource utilization and development," Which is still in the process of exploring specific methods in local governments and support staff assignment according to their abilities, as well as realize efficient human resource development.

Regarding the attitude toward viewing video materials according to ability, past research has identified the "time of video with the largest number of people leaving" based on Vimeo viewing history data and designed educational materials with this time in mind. Still, we could not analyze which parts of the video were practical or whether there was a relationship between the trend of the video itself and the attitude toward the course. However, we could not analyze from the perspective of which parts of the videos were practical and whether there was a relationship between the videos' trends and the students' attitudes. Therefore, through this study, we would like to gain knowledge on the design of courses according to ability.

2. Related Studies

2.1 DX Literacy Standards in Japan

The Digital Skill Standards (Ministry of Economy, Trade and Industry, December 21, 2022) were announced as the roles of personnel who promote DX and the skills they should acquire. It consists of two types of standards: the DX Literacy Standards, which are guidelines for all businesspeople to acquire basic knowledge, skills, and mindsets related to DX, and the DX Promotion Skill Standards, which are guidelines for companies to develop and hire personnel with expertise in promoting DX. However, the DX Literacy Standards do not mention local government officials.

For local government officials, the Ministry of Internal Affairs and Communications (MIC), which has jurisdiction over local governments, has prepared "Initiatives for Staff Development to Promote Municipal DX, June 3, 2022." Still, it remains vague guidelines for initiatives, does not refer to specific learning content such as DX literacy standards, and does not envision the creation of skill standards or the establishment of skill surveys targeting local government officials. Furthermore, it does not envisage the creation of skill standards or the establishment of a skill survey targeting local officials.

Therefore, a database of questions corresponding to the DX Literacy Standards will be created, and IRT will be used to evaluate the validity of the questions. However, since items such as "customer" and "competitive environment" cannot be applied as they are when

focusing on the work-relatedness of local officials, the DX Literacy Standards are redefined in terms of "citizen" and "social cooperation," respectively, and questions are created.
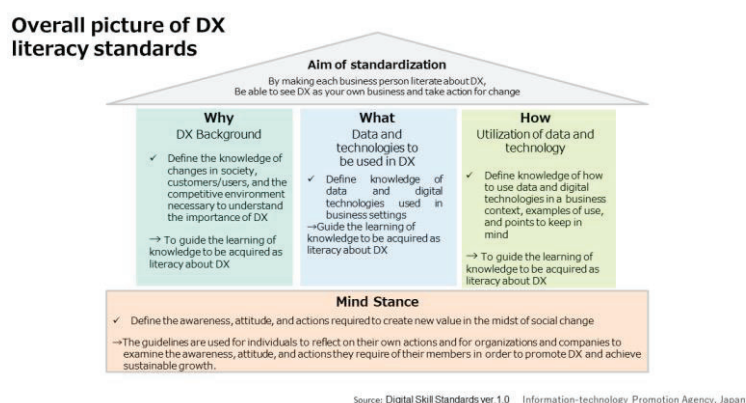

Figure 3 DX Literacy Standards

2.2 Creating a Skill Survey Environment Using IRT

Study items are presented in the DX Literacy Standards, but no examples of specific questions are given. Therefore, creating a surveyable question database for each area of DX Literacy is necessary. In Japan, the Information Technology Engineer Examination is a public examination that can be used as a reference for creating questions. The IT Passport Examination is the least difficult of these examinations. It covers so-called IT literacy, but the syllabus for the IT Passport Examination needs to cover the items required by the DX Literacy Standard.

Therefore, it is necessary to create questions that satisfy the items of the DX Literacy Standards while referring to the questions of the IT Passport Examination. The appropriateness of the questions in terms of content and difficulty will be analyzed using IRT, and inappropriate questions will be modified or eliminated. This test edition will be done by asking about 30 prior testers who have cooperated in formative evaluations in previous studies to answer the entire question database.

Once the appropriateness of the problem difficulty level has been confirmed to some extent, a large-scale skill survey will be conducted for all Hiroshima City employees. In this case, adaptive IRT will be employed.

For example, surveying a database of 200 questions by having the examinees answer all of them would require approximately 2 hours, and many examinees would "leave" the test. Therefore, to measure ability in 10 to 15 minutes, the test environment should be constructed as an adaptive test of the 2PL model, taking into consideration that "the question database contains enough questions with a wide variety of difficulty levels" (Yoshiko Tokusada, Takenori Sakumura, Naoki Tabuchi, Hideo Hirose,2019). Therefore, the test environment is constructed as an adaptive test of the 2PL model.

As a follow-up study to this research, we envision the development of a DX skills information dashboard that measures the distribution of DX skills among employees at the municipal level through a skill survey of local government employees nationwide and supports the planning of DX human resource allocation and training according to the results of the study. Therefore, Microsoft Azure SQL Database was selected as the database, considering the ability to utilize Azure Machine Learning as a data analysis tool (Figure 4).
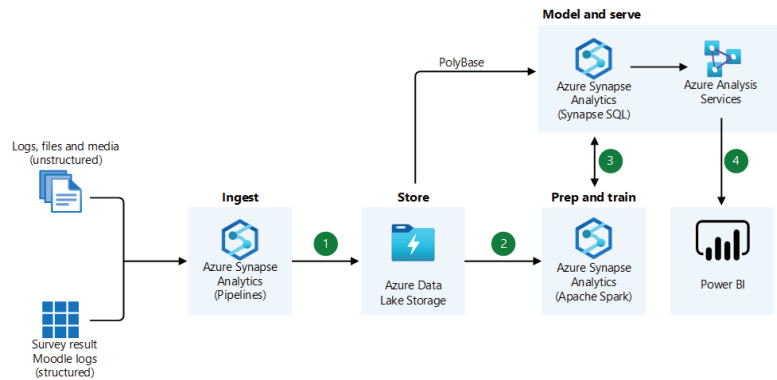
Figure 4 Azure Machine Learning Model

3. Perspectives on Analysis of Video Materials

3.1 Infrastructure to be constructed in this project

Moodle will be used for learning courses. However, the learning infrastructure will be improved from the following perspectives to realize more diverse data analysis. About the Study Course Moodle will continue to be used for learning courses.

Separation of test and course of the study as an extension module of Moodle, a previous study to realize IRT is "Development of Plug-in for Moodle: Adaptive Test Module based on Potential Rank Theory. " (Hiroshi AKIYAMA,2012)

However, with this method, it is necessary to respond appropriately to upgrades of the PHP used in Moodle and to promotions of Moodle itself. Furthermore, since there are plans to develop subsequent applications, such as aggregating the results of the skill survey for each local government, it was decided to create an entirely new system for the IRT test environment.

The video playback status should be able to be analyzed for each student. In addition, Azure Media Services will be used to analyze the videos to explore the relationship between their characteristics and the students' attitudes.

3.2 Focus on the video audience.

Have them take a skill survey and then take a course.

In doing so, we will focus on the relationship between the skills of each item measured in the skill survey and the data on the attendance at the educational materials created for each item, and whether there is a relationship between prior skills in each domain of the DX Literacy Standards and the attitude toward viewing the video materials corresponding to each field. For example, we will identify events such as skipping video playback in areas where prior skills are high and repeating video playback many times in areas where skills are low, as well as viewing attitudes such as the tendency for participants with higher overall skills to play videos faster.

3.3 Focus on the video itself.

Using Azure Video Indexer, the video material itself is analyzed. This function allows us to infer, for example, that specific keywords are repeatedly viewed, in which case it can be concluded that the explanation of those keywords is challenging to understand. Furthermore, if the speaker leaves after a certain period of silence, it can be inferred that there is a silent period during which the learner is bored. In this case, there is a quiet period during which the participants are bored. Table 4 shows the analysis items of Azure Video Indexer, which is planned to be used in this project.

Table 4. Azure Video Indexer Analysis Items

statistics		labels	
Name	**Description**	**Name**	**Description**
CorrespondenceCount	Number of correspondences in the video.	id	The label ID.
WordCount	The number of words per speaker.	name	The label name (for example, 'Computer', 'TV').
SpeakerNumberOfFragments	The amount of fragments the speaker has in a video.	language	The label name language (when translated). BCP-47
SpeakerLongestMonolog	The speaker's longest monolog. If the speaker has silences inside the monolog it's included. Silence at the beginning and the end of the monolog is removed.	instances	A list of time ranges where this label appeared (a label can appear multiple times). Each instance has a confidence field.
SpeakerTalkToListenRatio	The calculation is based on the time spent on the speaker's monolog (without the silence in between) divided by the total time of the video. The time is rounded to the third decimal point.		

3.4 Their relevance

By combining video playback data per viewer unit with the analysis results of Azure Video Indexer, we can, for example, extract keywords that are difficult for students with low scores in specific areas in the DX Literacy Standards, relate playback speed to ability before taking the course, and confirm the effectiveness of improving teaching materials based on these data. The data will be used to verify the effectiveness of enhancing teaching materials.

User	Topic	Level 1	Level 2	Level 3	Attitude	KeyWord
A	AI				Very Fast,Skip	—
B	AI				Quit	—
C	AI				Slow	Deep learning,Neural network
A	Network				Fast	subnet mask
B	Network				Normal	subnet mask,IPv6
C	Network				Slow	subnet mask,IPv6,DNS
A	Security				Very Fast	—
B	Security				Slow	Availability,AES,Public-key cryptography
C	Security				Fast	Availability,AES

Figure 5 Expected Results

4. Future Prospects

4.1 Standardization of DX Literacy Skill Survey in Local Governments

The provision of skill surveys to measure DX literacy standards needs to be improved, partly because it is still early days since the digital skills standards were published.

Furthermore, since this study is the only one that reads the DX literacy standard regarding work relevance for local governments, it is expected to serve as a skill survey standard for local government employees.

4.2 Improvement of teaching materials

Continuously improve teaching materials by analyzing data to understand the responses of "voiceless participants," which could not be ascertained through formative evaluation.

4.3 Creating an Individualized Optimization Environment

First, efforts will be made to improve the completeness of the skill survey and e-learning materials. Indicators include: the results of the IRT analysis of the skill survey question database should be appropriate; the quality of the video materials for each of the DX Literacy Standards should exceed 90% as "easy to catch on"; and the quality of the video materials should exceed 100% as "easy to learn." A 100% pass rate on the post-learning test within three attempts. Achieve these conditions.

The goal is to construct a system that can propose the optimal order of participation for participants to master all DX literacy standards by measuring changes in their willingness to learn when they take the courses in the order of their strengths and weaknesses in the skill survey results, and the difference in the number of days it takes them to complete each course item before starting the next.

238

	TOPIC	Survey	Course		Survey	Course		Survey
Why	Social changes	○	Y-1		—	Y-1		—
	Changes in customer value	○	Y-2		—	Y-2		—
	Changes in the competitive environment	×	Y-3		○	Y-3		—
What	Data in society	○	T-1		—	T-1		—
	Reading and explaining data	○	T-2		—	T-2		—
	Dealing with data	○	T-3		—	T-3		—
	Determined by data.	×	T-4		○	T-4		—
	AI	×	T-5		○	T-5		—
	Cloud	×	T-6		○	T-6		—
	Hardware and software	×	T-7		○	T-7		—
	Networks	×	T-8		×	T-8		○
HOW	Examples of the use of data and digital technologies	○	W-1		—	W-1		—
	Tool use	×	W-2		○	W-2		—
	Security	×	W-3		○	W-3		—
	Morale	×	W-4		○	W-4		—
	Compliance	×	W-5		○	W-5		—

Analysis of survey responses and course attendance data.

Discover the best sequence of courses to take.

Figure 6 Optimization through skill surveys

Acknowledgments

References

Minoru AKIYAMA. (2012) Development of plug-ins for Moodle -Adaptive testing module based on latent rank theory.

Mitsuko TOKUSADA., Takenori SAKUMURA., Naoki TABUCHI., Hideo HIROSE. (2015). Considerations on the Optimal Number of Questions in Adaptive IRT Testing.

Marcella Milana., Sue Webb., John Holford., Richard Waller., Peter Jarvis. (2017) The Palgrave International Handbook on Adult and Lifelong Education and Learning

Makoto Takahashi. (2021) Introduction to IRT Item Response Theory: How to make a quality test from the basics of statistics.

Kentaro Kato., Takeshi Yamada., Kazumitsu Kawabata. (2023) Item Response Theory with R

Yuhiko Mitsunaga., Akiko Nishida. (2022) What are tests for? - Considering the Admissions System from Item Response Theory.

Microsoft Corporation. (2023). Analyze video and audio files with Azure Media Services. https://learn.microsoft.com/en-us/azure/media-services/latest/analyze-video-audio-files-concept

Microsoft Corporation. (2023). Dynamic packaging in Media Services v3. https://learn.microsoft.com/en-us/azure/media-services/latest/encode-dynamic-packaging-concept

Ministry of Economy, Trade and Industry. (2022). DX literacy standard.

Ministry of Internal Affairs and Communications. (2022). Initiatives for Staff Development to Promote Municipal DX.

Hiroaki Ogata., Etsuhiro Eguchi. (2023) Learning Analytics to Transform Learning

Reza Rad. (2023) Pro Power BI Architecture: Development, Deployment, Sharing, and Security for Microsoft Power BI Solutions

Louis Davidson. (2023) Practical Graph Structures in SQL Server and Azure SQL: Enabling Deeper Insights Using Highly Connected Data

Chris Seferlis., Christopher Nellis. (2023) Practical Guide to Azure Cognitive Services: Leverage the power of Azure OpenAI to optimize operations, reduce costs, and deliver cutting-edge AI solutions

Maxwell Everstone. (2023) Azure Data Factory: A Comprehensive Guide

Mervin Weiden. (2023) Mastering SQL Server: Streamlining SQL Server Always on Availability Group Deployment In Azure