A Survey of Student's Perception for Information Literacy

Ayako OHSHIRO[†*]

[†]Department of Business Administration Okinawa International University
Ginowan City, 0kinawa 901-2701, Japan
*Correspondence

Summary
It is important for educators to understand students' awareness of information literacy, in order to provide them with a good information literacy education that will help them function in an advanced information society. This study administered a questionnaire survey to measure students' recognition of the importance of various information literacy categories in different life situations. In this paper, we report on the data analysis results by using text mining and data statistics, and consider approaches for training on information literacy education training that is focused on students' awareness.

Key words:
Information literacy education, Text mining, Data analysis

1. Introduction

Given the popularization of equipment such as smartphones and tablets, information technology tools have become a part of life. Furthermore, because most college lectures have migrated to online platforms in recent times in response to novel coronavirus pandemic, the internet environment can be said to be necessary for education. In addition, precipitated by health safety measures, the need for personnel who can work remotely has been growing, not only in the field of education, but also in society at large. These circumstances have revealed that information literacy levels are insufficient to enable individuals to acquire the information they need on a daily basis in modern society; it is therefore necessary to provide information education. Although more than ten years have passed since information has been a compulsory subject, it has a short history compared to other subjects, hence the development of an appropriate training method for this area is still underway. Efforts include a review report on a college's information educational curriculum, based on the high school registration pattern for the information subject [1]; the collection of first-year college students' ideas about information education [2]; the consideration of teaching style, based on students' computer environment [3]; the study of information literacy's ability to influence student's performance in other subject areas [4], and the relationship between various student characteristics [5]. Moreover, numerous investigations such as the minimal information literacy level required for each specialized field [6], information skills as common subjects that are required by each university department [7], and an effectiveness measurement of instructional materials based on the traditional education model [8] have been reported. Information education content has widely included computer knowledge, information morals, and programming, and it is thought that the instructional points will vary by educator.

However, many educators recognize that Microsoft Excel training is particularly significant [1], because most students have weaknesses in terms of their ability to use the program [3], despite the fact that so many situations, such as graduation research and subsequent social life, require data handling using Excel skills, including filtering, functions, and so on. Globally, spreadsheet skills are required for a certain level of work, regardless of occupation [9]; this, coupled with the fact that media literacy and information literacy are basic competencies in most curricula, means that college students need to acquire these capabilities before becoming functional members of society [10]. Therefore, the purpose of university information education in the context of information society is to convert students' weaknesses in terms of using Excel into recognitions as important via training. In addition, the sharing of information literacy awareness between educators and students may have a positive effect on training. For that reason, it is necessary for educators to accurately determine student's information literacy recognition, background, and potential. The recent years have seen an increase in the number of studies on the extraction of potential features that cannot not be extract from quantity data by text mining student's free descriptions [11] and in the number of studies on generating educational model with extracted features [12].

In this research, an investigation of college students' recognition including their backgrounds, with regard to each information literacy category, was conducted. In this paper, we report and consider the result of applying text mining and analysis to the questionnaire data, which consist of the allocation and document forms.

Manuscript received September 5, 2020
Manuscript revised September 20, 2020

DOI: 10.22937/IJCSNS.2020.20.09.7

2. The process of surveying student perception

In this chapter, we describe the process of surveying student perception, from investigation to data analysis. In the author's information literacy class at several collages, the syllabus is delivered based on a text that prepares students for the ICT-proficiency assessment. Since the text covers a wide range of content, such as computer knowledge, information morals, Excel, and Word, it is treated as a measure of information ability at many universities [13]. First, with regard to students' scoring on recognition, five categories were defined as follows, with reference to the ICT-proficiency assessment: C_1 : Computer knowledge, C_2 :Information and communication network, C_3 :Information morals and security, C_4 :Word, and C_5 :Excel. This study assumed two situations, namely "college student life" and "working life after graduation"; the description form concerned recognition of the selected category. This study also adopted the assumption that it is characteristic among college students for there to be some difference between first-year students who are in the early post-high-school period and second-year students who have had time to adjust to college life and take an interest in social life. We confirmed the differences by year of study by classifying the students into 'lower' and 'upper' academic grades. Moreover, we assumed a potential element with high-frequency words, which were extracted by applying morphological analysis to the description data. Second, each category feature is compared with the statistical value. As will be described later, because the amount of collected data depended on year of study, the coefficient of variation was utilized for measuring variability.

From the above, we describe the process of administering the survey and conducting a comparison study to assess college students' recognition of the importance of information literacy as follows.

Step 1: Request that students evaluate "categories that are felt to be useful in each situation, namely student life (***LS***) and society life (***LSO***)" in the selective questionnaire distributed that was among the students sample for each information category factor, such as C_i, making the total value 1.0.

Step 2: Obtain free description data for each category, as evaluated by the students in the Step 1.

Step 3: Apply the result of morphological analysis to the free description data collected in Step 2 and extract "noun" and "verb" as *Terms*, and deleting "number," "independence," "non-autonomy," and "suffix."

Step 4: For the result obtained in Step 3, extract words that appear frequently are extracted.

Step 5: Calculate the statistical value of each category, such as *average* (***A***), *median* (***M***), and *coefficient of variation* (***CV***).

Step 6: Confirm the difference by academic grade (year of study) by classifying them into lower grades, ***GL*** and the upper grade, ***GU***, using analysis of variance.

3. Collected data

The survey was carried out between 2018 and 2019 in four classes at two universities. In this investigation, questionnaire data were collected from 85 upper-grade students and 150 lower-grade students.

4. Analysis results and discussion

In this chapter, we discuss students' recognition of the importance of information literacy from many perspectives, using analysis result. First, the statistical values of each category and the high-frequency key terms in the cases of student life and social life are shown in Tables 1 and 2 for the lower grade and Tables 3 and 4 for the upper grade. Terms are sorted in descending order of frequency of appearance and written as "Frequency value : *Term*" in the tables.

Table 1. <Lower grade・Student life>Statistical values and high-frequency terms for each category

	A	***M***	***CV***	***Term***
C_1	**0.3448**	0.300	0.5620	15: Computer 12: Knowledge 5: Necessary 4: Basic 3: USB 2: Shortcut key 2: Data backup
C_2	0.2741	0.200	**0.6536**	2: Information 1: Mail 1: Study 1: Search 1: Report 1: SNS
C_3	0.3127	0.300	0.5254	11: Information 7: Internet 7: Security 4: Myself 2: Manners
C_4	**0.5184**	0.500	**0.4687**	69: Report 11: Write 15: Lecture 5: Most necessary 3: Knowledge 3: Figure 2: Seminar
C_5	**0.3328**	0.300	0.4823	10: Table/Report 8: Lecture 9: Graph 4: PowerPoint 3: Presentation 3: Calculation 5: Functions 2: Knowledge

Table 2. <Lower grade・Society life>Statistical values and high-frequency terms for each category

	A	*M*	*CV*	*Term*
C_1	**0.3714**	0.400	0.5403	28: Computer 21: Knowledge 6: Information 6: Work 8: Information 4: Network
C_2	0.2742	0.200	**0.6598**	4: Internet 2: Communication 2: Work 2: Knowledge 5: Work 5: Society
C_3	0.2986	0.250	0.5859	5: Knowledge 4: Moral 4: Virus 4: Security 2: Trouble 38: Report/Document
C_4	**0.3494**	0.300	0.5490	14: Work 5: Presentation 9: Summarize 7: Getting a job 4: Notion 3: Report 3: Right/Invasion 2: Project
C_5	**0.3364**	0.300	0.5078	2: Routine 2: Trouble 18: Summarize 16: Work/Company 15: Documents 11: Data 11: Table 10: Preparation 6: Presentation 5: Calculation/Graph 3: Knowledge 2: Efficiency

Table 3. <Upper grade・Student life>Statistical values and high-frequency terms for each category

	A	*M*	*CV*	*Term*
C_1	**0.2946**	0.200	0.7729	7: Lecture 8: Computer 4: Knowledge/Report 2: PowerPoint/ Treatment/Project /Information /Research Paper/Write
C_2	0.2671	0.200	0.732	8: Computer 7: Lecture 4: Knowledge/Report 2: PowerPoint/Treatment /Project 1: Information/Write /Research Paper
C_3	0.2185	0.200	0.4563	2: Computer/Rules /Damaging/Application 1: Software/Rule/Item **/Unconscious** /Information/Myself /Fraudulence /Invasion/Manners /Keep in mind
C_4	**0.4901**	**0.500**	0.4162	**46: Report/Lecture** **8: Submission** **/Documents** **5: Write 4: Summarize** **3: Presentation** **2: Research Paper** 1: Header/Footer
C_5	**0.377**	**0.400**	0.3982	**7: Report** **6: Figure/Summarize** **5: Graph/Statistics/List** **5: Seminar** **4: Questionnaire** **/Aggregate/Calculation** **/Document/Information** **3: Lecture 2: PowerPoint** 1: Investigation/Research /Paper/Numerical value /Copy/Lump

Table 4. <Upper grade・Society life>Statistical value and high-frequency terms for each category

	A	*M*	*CV*	*Term*
C_1	**0.3118**	0.225	0.7112	22: Computer 14: Knowledge 5: Information 4: Basic 3: Data 2: Important/Getting a job /The least/Database 1: Skill 6: Information
C_2	0.2234	0.2	0.6854	1: Contact/Future/ World/Knowledge/Addition /Come through/Obtain /Nowadays/Collection /Network 16: Information 9: Follow
C_3	0.2759	0.2	0.5124	8: Moral/Security 7: Knowledge 5: Important 3: Deal 2: Myself/Individual /The least/Leakage 1: Prevent/Rules
C_4	**0.2958**	0.25	0.6336	12: Documents 5: Report 2: Basic 10: Calculation/Statistics /Numerical value/Aggregate /Information
C_5	**0.3184**	0.3	0.5745	9: Data 4: Graph 3: Getting a job/Documents /Absolutely 2: Analysis/Figure/Increase /Management/Summarize /Fixing

As shown in Table.1, C_4 has the highest average and median, and terms such as "report" "lecture," and "seminar" were extracted as key terms. In addition, C_4 has the smallest coefficient of variation. This means that lower-grade students place great importance on Word skills because it is heavily used in many classes in student life. However, this result could not be clarified from the educator's position. C_1 has the highest average after Word skills and "USB handling," "shortcut key," and "backup data" were extracted as key terms. This result suggests that computer knowledge assists students in daily computer-use life. On

the other hand, C_1 has the highest average and median, and followed by C_4. "Work" was extracted as a key term with regard to computer knowledge, and words such as "document," "work," and "summarize" were extracted with respect to Word skills, as shown in Table 2. This suggests that lower-grade students recognize that computer knowledge will assist them in society. Although C_3 has the smallest average, extracted terms such as "unconscious/invasion/fraudulence/manners" indicate that students feel that is necessary to learn about information morals and security. Meanwhile, C_2 is weakly recognized in both student and society life. It is thought that students have a professional image and are not familiar with network knowledge. Moreover, C_2 has the largest coefficient of variation in both Tables 1 and 2, implying that students' recognition of the importance of information and communication networks is too varied for any solid conclusions to be drawn at this time.

Next, we will focus on recognition in the upper grade. As shown in Table 3, C_4 has the highest average and median, and "report," "lecture," "submission," and "graduation thesis" were extracted as key terms, thus bearing some similarity to lower-grade students. Moreover, C_5 has the second highest average score after Word skills, and "report," "graph," and "aggregate" were extracted as key terms with respect to data collection or data management. Notably, median scores are biased to Excel (0.5) and Word (0.4), indicating that upper-grade students tend to have strong recognition of the importance of acquiring these skills. Table 4 shows that C_5 has the largest median and average scores in society life, C_1 is second, and C_4 is ranked third. The words "statistics" and "numerical value" are newly-extracted with respect to Excel skills, suggesting that upper-grade students recognize that they often handle data in society life. In addition, "the least/the element" was extracted in the case of computer knowledge, and "report/the element/document" was extracted in the case of Word skills. Those tend to be recognized as having minimal importance in terms of the acquisition of skills that will be useful in society.

The above findings show that both the upper and lower grades recognize "computer knowledge" as common knowledge, and with lower-grade students exhibiting strong recognition of "Word-skill" and becoming more aware of "Excel-skill" and "computer knowledge" as they advance to the higher grades.

Since it has been shown that there is a possibility of differences in the recognition for the importance of information literacy depends on year of study, Analysis of variance (ANOVA) was used to confirm this as shown Table 5.

Table 5.The *p*-value in ANOVA for each information category by academic grade/years of study

	C_1	C_2	C_3	C_4	C_5
LS	0.2109	0.8192	0.2100	0.4198	**0.0008**
LSO	0.1914	0.9391	**0.1182**	0.3183	0.7226

Since the *p*-value of C_5 for *LS* is significantly small, it means that recognition of the importance of Excel for student life differs significantly by academic grade-year of study. This suggests that above-second-year students have more opportunities to analyze big data for practical subjects and/ or graduate research. Similarly, the *p*-value of C_3 for *LSO* is significantly small, implying that recognition of the importance of information morals for society life also differs by year of study. We can conclude that above-second-year students make some contact and thus become relatively familiar with society and its skill demands while job hunting.

In order to correlate recognition, the coefficient of correlation between each information category (C_1 - C_5) for *GL* or *GU* and student life (*LS*) or society life (*LSO*), are shown in Table.6.

Table 6. The coefficient of correlation between information categories by academic grade/year of study

GL		C_1	C_2	C_3	C_4
LS	C_2	0.10941			
	C_3	0.00323	0.1179		
	C_4	**-0.4655**	-0.3192	-0.3361	
	C_5	-0.1896	-0.1322	0.1230	-0.1740
LSO	C_2	-0.0377			
	C_3	-0.064	0.0345		
	C_4	-0.3448	-0.3123	-0.3395	
	C_5	-0.3481	0.2134	-0.3412	0.2761
GU		C_1	C_2	C_3	C_4
LS	C_2	0.0866			
	C_3	0.0161	0.3217		
	C_4	**-0.4568**	-0.3888	-0.2222	
	C_5	-0.2742	-0.3151	0.1855	0.0554
LSO	C_2	0.0090			
	C_3	-0.0423	-0.1049		
	C_4	-0.3352	-0.2782	-0.2066	
	C_5	-0.3273	-0.2071	-0.2960	0.1348

A slight negative correlation between C_1 and C_4 was found in the cases of *GL* and *GU,* suggesting that recognition under those circumstances is biased compared to other combinations. However, the bias decreased by about 0.1 in *LSO* compared to *LS,* and as mentioned, students recognize the importance of Word skills for reporting during their lower years of study and come to recognize the wide utility of information literacy for society life as they advanced to the upper years of study.

5. Conclusion

With the goal of providing good practices for information literacy education, a survey was conducted to determine students' recognition of the importance of several related subareas. Specifically, the questionnaire items pertained to the recognition of information categories related to student life and society life. Data analysis results showed that students' recognition of information category changes with the environmental changes that are concomitant with advancing through years of academic study. Based on insights obtained through this perception research, we hope to train college students so that they are not only capable of responding to shifts in college life, but are also highly adaptable to information society after college graduation.

References

[1] Tomio Kuroda, Takeo Okazaki, "Agenda for the Information Study, a newly introduced subject of school—Verifying the new subject Information in senior high school:Registration pattern for new subjects Information and their average learning level of the senior high school in Okinawa Prefecture—", pp.30-pp.35 VOL.21, Computer&Education 2006

[2] MASANORI TAKAGI, MITSUHIRO SETOYAMA, "Investigation of the MS-Office Operating Skills of Freshman Using Computer-adaptive Testing", pp.215-pp.220, Information Processing Society of Japan, 2018

[3] OHTA Nobuhiro, "A Study on Freshmen's Computer Literacy and Office Software Skills", p.3-p.14, vol. 29 No.2, Japan Society of Educational Information, 2014

[4] Scott Lanning, Jill Mallek, "Factors Influencing Information Literacy Competency of College Students", *The Journal of Academic Librarianshi, p443-p450,* Vol. 43, Issue 5 2017

[5] Ove Edvard Hatlevika, Inger Throndsen, Massimo Loi Greta B.Gudmundsdottir "Students' ICT self-efficacy and computer and information literacy: Determinants and relationships", Vol.118, p.107-p.119, 2018

[6] Kei TAKAHASHI, Kenji SAITOH, Mitsuo MITA,Kazuo KATO,Takeshi KASHIMOTO, Tokio NAWA,Saburo HORIUCHI, "Development of Computer Literacy Based on Education and Practical Exercise" Medical Education (Japan) 2000, 31(3): 171-179

[7] Tomio Kuroda, Takeo Okazaki, "About the contents of a class of Information Processing Education as a General Education and that goal for arrival.", no.10 p.118 -124, University Education Center Bulletin, 2006

[8] Shirakawa Yuzo, Kamiya Yoshihiro, Nakajima Koji, "Measurement of Effectiveness of Applying Instructional Design in a Teaching Material of Information Literacy", Vol.30, No.3, p.61-70, Japan Society of Educational Information 2015

[9] Sam K. Formby, B. Dawn Medlin, Virginia Ellington, "Microsoft Excel®: Is It An Important Job Skill for College Graduates?", Information Systems Education Journal (ISEDJ) Vol 15, No. 3, 2017

[10] S.Ingvaldsen, "Chapter 4 - The School Library in Media and Information Literacy Education", p. 51-p.65, Media and Information Literacy in Higher Education, 2017

[11] Senji YAMAZAKI, Junichi KAKEGAWA, Hisashi OGAWA, Naoki KATO, Ritsuko OKIDO, Koichiro MORIHIRO, "Comparison of Text−Mining Methods for Extracting Feature Words from Elementary School Report Card Comments", vol. 31, No.2, p.37-p.48, Japan Society of Educational Information 2015

[12] Ayako OHSHIRO, Shinichiro UEDA, "Factors extraction for Clinical Research educational model and their Relation", IJCSNS International Journal of Computer Science and Network Security, VOL.19 No.3, pp.87 -91, March 2019

[13] Hirokazu Murakami, Hatsuo Yamasaki, Yukimasa Terada, "An inspection of skills of a computer and a literacy education for freshumen", vol. 108, no. 88, ET2008-6, pp. 1-8, 2008

Ayako Ohshiro received the B.S. and M.S. degrees from the University of the Ryukyus in 2009 and 2011, respectively. She took Ph.D. from University of Ryukyus in 2017 and she had been an assistant professor at University of the Ryukyus from 2015 to 2020. She has been a lecturer at Okinawa International University since 2020. Her research interests are data analysis in the field of intelligent information engineering.