MS. Dissertation in Engineering

A Study on Online Education Users' Information privacy perception and Decision-making process

온라인 교육 이용자의 정보 프라이버시 인식과
의사결정 과정에 관한 연구

August 2023

Graduate School of Seoul National University

Technology Management, Economics, and Policy Program

Jieon Park

A Study on Online Education Users' Information privacy perception and Decision-making process

지도교수 황준석

이 논문을 공학석사학위 논문으로 제출함
2023 년 8 월

서울대학교 대학원
협동과정 기술경영경제정책 전공
박지언

박지언의 공학석사학위 논문을 인준함
2023 년 8 월

위 원 장 **Jörn ALTMANN** (인)

부위원장 황 준 석 (인)

위 원 윤 현 영 (인)

Abstract

A Study on Online Education Users' Information privacy perception and Decision-making process

Jieon Park

Technology Management, Economics, and Policy Program

The Graduate School

Seoul National University

Online education is a type of distance education, which enables education anytime, anywhere. With the development of digital technology and other more advanced technologies such as artificial intelligence and learning analytics, gathering and analyzing data in the education sector is growing. However, not many studies have investigated privacy issues in the context of education. Therefore, in this study, the educational service users' perception of privacy and their according behavior will be investigated in an effort to analyze the privacy issues in the educational field.

The main theories are privacy paradox, and privacy calculus. The privacy paradox explains a situation where individuals decide to give away their

information even though they are concerned about disclosing their information. Privacy calculus theory is one of the theories that explain this paradoxical phenomenon, where individuals economically and rationally compare the benefit and risk of disclosing information before deciding what to do. Survey and Structural Equation Modeling (SEM) were conducted for analysis and the hypothesis testing result showed that there are privacy concern, and privacy paradox in the education sector also, with privacy calculus theory explaining why individuals decide to reveal their information.

Keywords: online education, privacy concern, privacy risk, privacy benefit, privacy calculus

Student Number: 2019-26284

Contents

List of Tables

List of Figures

Chapter 1. Introduction

The advent of the internet has digitalized numerous fields such as commerce, governance, media, and healthcare. The field of education is not an exception. In the case of Korea, technologies such as Information and Communication Technologies (ICT), online learning platforms and learning management system (LMS) enabled digitalization of education. The characteristic of these technologies is that they collect, analyze, and utilize innumerable data generated from educational service users. Therefore, preparing appropriate data utilization and personal information protection method by studying users' perception and their actual behavior of data usage is becoming more important than ever. In other words, studying information privacy in the context of online education is becoming crucial.

According to studies related to information privacy, although digitalized service users are concerned about their information disclosure, they disclose their information to use the service. This behavioral tendency is known as 'privacy paradox'. Privacy paradox explains the situation where the internet service users disclose their information to service providers in order to use the service, even though they are concerned about revealing their information. This phenomenon was first examined by Brown in the early 21st century, by conducting in-depth interview of online shoppers. After that, a number of researchers scrutinized this

phenomenon usually by quantitative study, especially survey, in various contexts. The contexts include online shopping, social media, website, healthcare and governance. Privacy paradox was usually studied in the context of online shopping and social media.

One of the renowned theories explaining the privacy paradox phenomenon is ‘privacy calculus theory’. This theory postulates that humans are rational beings, as they compare the risk and benefits of information disclosure. According to this theory, individuals decide to giveaway their information as they believe the benefit of doing so outweighs the risk. Examples of other theories that explain privacy paradox are ‘bounded rationality’, ‘heuristics’, ‘social influence’, ‘the risk and trust model’, ‘quantum theory’ and so on.

However, although there are many studies proving and explaining the paradoxical behavior of online service users, only a few studies scrutinized this in the context of online education. As mentioned before, digital technology is being widely used in the education sector. The aftermath of COVID-19 has accelerated the adoption and diffusion of digital technologies in educational field. Moreover, contrary to the popular belief that the younger generation is relatively less sensitive to privacy, sensitivity toward information privacy changes depending on the context of the situation or social factors, underlining the importance of studying privacy paradox in online education sector.

Therefore, this study aims to analyze online education users’ awareness,

decision making process and behavioral intention on the use of personal information. As the importance of personal information utilization and protection in the field of education is emerging, the research findings will help exploring ways to use information appropriately in the online education era.

This research is organized as follows: Chapter 2 is consisted of literature review of online education, information privacy, privacy paradox and trust on institution. Next, chapter 3 handles research methodology, research model, definition of variables and introduces hypotheses. Chapter 4 presents result of this research along with reliability, validity test and Confirmatory Factor Analysis (CFA). Last but not least, chapter 5 provides summary of research, and goes through implications and limitations of this research.

Chapter 2. Literature Review

2.1 Online education

2.1.1 The concept of online education

Online education, also interchangeably termed as e-learning, utilizes the Internet technology to deliver learning activities and resources of various types and ranges to improve knowledge and performance (Rosenberg, 2002; Badrul Khan, 2005). Specifically, the terminology e-learning emphasizes the use of the technology as the prefix e- stands for electronic (Moore & Greg Kearsley, 2011), whereas online learning emphasizes the educational environment. This point of view is found in the definition of Nada Dabbagh & Brenda Bannan-Ritland (2005), as they define online learning as an open and distributed learning environment enabled by Internet and Web-based technologies.

Although there are many different definitions of online education, the essence of online education is categorized educationally and technologically. The educational essence is that online education's characteristics are interactive, open-ended and flexible. The technological essence is that online education uses Information and Communication Technologies (ICT) (한국교육공학회 et al., 2016).

Moreover, online education provides Learning Management System (LMS) as

educational environment (한국교육공학회 et al., 2016). LMS is a management system for education which provides primary functions such as record management, design and provision of classes, and evaluation; secondary functions of LMS are communication, management of student and faculty data (W. R. Watson et al., 2007; Rita C. Richey, 2013). In other words, LMS is a system that manages the entire process of online education. The representative LMS providers are Moodle, Blackboard and Canvas.

Another explanation for the concept of online education is that it is a type of distance education, that arose with the development of technology. Distance education emerged and developed with the effort to provide education, anytime, anywhere, regardless of time and place. The development of distance education can be divided into three generations. However, the advent of a new generation does not mean the replacement of the prior generation, rather it is a coexistence.

The first-generation distance education, the earliest form of distance education that started from the 19th century, is conducted through print media and postal systems. The second-generation distance education is conducted through mass media such as radio, television, and telephone. This type of distance education was widely used from the 1960s to 1990s, before the internet and computer were introduced in the education field. Last but not least, the third-generation distance education, which continues to this day, has been enabled by the development of Information Communication Technology. With the help of computers, computer-

mediated communication, telecommunications satellites, Internet and so on, the third-generation distance education has made real time, two-way interaction possible (조은순 et al., 2012). Online education falls into the third-generation distance education. Table 1 summarizes the media or technologies used in different types of education.

Table 1. Media / technologies used in education

Category	Media / Technologies
First-generation distance education	Print media, postal systems
Second-generation distance education	Radio, television, telephone, broadcasting systems
Third-generation distance education	Computers, computer-mediated communication, telecommunications satellites, Internet

2.1.2 Data collection and utilization in online education

In order to deliver and conduct online education efficiently, an infrastructure is needed. As mentioned previously, the Learning management system (LMS) is an infrastructure that governs the entire process of learning. In detail, the LMS "delivers and manages instructional content, identifies and assesses individual and organizational learning goals, tracks the progress, and collects and presents data for supervising the learning process of an organization as a whole" (Szabo, 2002;

William R. Watson & Watson, 2007). Not only the LMS delivers content and facilitates learning but also handles course registration and administration, skills gap analysis, tracking and reporting (Gilhooly, 2001). As can be seen from the functions of the LMS, the LMS collects and uses various data. The data

As collection of utilization of data in online services is an important issue related to privacy, the types of data collected from the LMS is presented by educational institutions' LMS, and LMS providers, as part of their privacy policies. Table 1 provides the types of data collected from LMS.

Table 2. Types of data collected from LMS

Category	Types of data
Administrative data	Civil status, identity, identification data, images etc.
Identifiers	Name, customer number, address, phone number, email address, date of birth, resident registration number etc.
Personal life data	Lifestyle, family situation, etc.
Economic and financial information	Income, financial situation, tax situation, etc.
Connection data	IP address, logs, cookie etc.
Educational data	Assessed coursework, exam scripts etc.
Records of educational attainment	Results of exams, assessments, qualifications awarded etc.

Location data	Travel, GPS data, GSM, etc.

Source: Retrieved and revised from Moodle website, Data Privacy section

2.1.3 Privacy issues in the context of online education

Privacy is a demand to be free from any surveillance and obstruction by another person, institution, or country (Kenneth C. Laudon & Jane P. Laudon, 2016). Furthermore, information privacy is a right related to the collection, storage, processing, and dissemination of personal information (Kokolakis, 2017). In this research, the terminology "privacy" specifically refers to the information privacy.

The classic concept of privacy was related to physical space (territorial privacy), and personal right (privacy of a person). The necessity to extend the scope of privacy to information privacy arose (H. Jeff Simth et al., 2011), as the development of Information Communication Technology started to gather personal and organizational information. Information privacy is increasingly emphasized as the development of Information Communication Technology gathers more detailed information, sometimes without consent or inevitably forcing information owners to allow gathering. Incidents such as information leakage to third party due to improper data management or hacking and abuse & misuse of personal, organizational data also emphasizes the significance of

information privacy.

Likewise, with the development of ICT, the education sector also started to collect and utilize data from the educational service recipients. Similar incidents related to information privacy issues happen in the education sector. For example, case of private equity firm's M&A of Canvas, one of the most largest LMS, for $ 2billion occurred in 2019 ((Jones et al., 2020), 2019; Jones et al., 2020). As the CEO of Instructure, the former owner company of Canvas, stated that the company has "the most comprehensive database on the educational experience in the globe.. enabling the development of those algorithms and predictive models", many inferred that the access to student data contributed to Instructure's financial value (Young, 2020; Jones et al., 2020). Although the private equity company and Instructure asserted that they would not share or sell student data, the acquisition stroked societal displeasure and concern (Young, 2020; Jones et al., 2020).

Moreover, investigating and preparing for privacy issues in education is becoming more important than ever as technologies that handle a massive amount of data are now being applicated in the education sector. The examples of these technologies are big data analytics, data mining and artificial intelligence. And the exemplary field being researched is learning analytics. According to Siemens et al. (2011), "Learning analytics (LA) is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs". In other words,

learning analytics requires more vast and diverse data compared to the LMS, as the examples of data gathered are interaction behavior log data, physiological data (heart rate, gaze).

However, in the education sector, not many researches are conducted on information privacy issues in education. Though a number of researchers mention the importance of privacy, privacy itself rarely becomes a research topic. Most studies that hold privacy as their research topic in online education focus on theoretical, ethical discussions [1] or present only descriptive statics. These approaches are essential in developing theoretical, ethical, legal, and policy-level foundations. However, for further theoretical and practical advancement, empirical study is also needed by investigating the actual perception and behavior of online education participants.

2.2 Privacy calculus theory

2.2.1 The Concept of privacy calculus theory

Privacy calculus theory is one of the most renowned approach in explaining online service users' information disclosure behavior. The theory postulates that online service users are economically rational beings, performing a comparison

[1] Related researches are presented as follows:
Education, Technology, and Individual Privacy, 1978
Who Is Reading Whom Now: Privacy in Education from Books to MOOCs, 2015
Learner Privacy in MOOCs and Virtual Education, 2018

between the expected risk and the potential benefit of information disclosure (Dinev & Hart, 2006; Xu et al., 2009; Li, 2012). Privacy calculus theory has two fundamental concepts. One of them is the assumption that human beings are rational, and the other is economical comparative calculation model. In other words, the rational humans compare the risk and benefits of information disclosure and trade their information if they decide the benefit is greater than the risk.

In this theory, the comparison of risk and benefit is crucial, as many online services' policies require users' information to use their services. Some may be reluctant to disclose their information but decide to do so, as they need or want to use the service. Therefore, the result of the privacy calculus decides whether the individual user will disclose their information or not. If individuals conclude that benefit is greater than loss, they decide to disclose personal information and use the service. If risk is larger than perceived benefit, the user will not disclose their information. In other words, the privacy calculus theory is a model that explains online service users' decision-making process.

2.2.2 Previous studies of privacy paradox

Privacy paradox phenomenon was usually studied mostly in the e-commerce, e-government, social media and healthcare sector (Spyros Kokolakis, 2017). The

term ‘privacy paradox’ was first shown in research related to internet use, especially in the context of online shopping (Brown, 2001). By conducting in-depth qualitative interview, Brown discovered that although online shoppers had high concern in relation to their privacy, it did not interfere them from giving out their information online for advantages such as loyalty cards. In other words, even if they felt their privacy might be infringed, they exposed their information if they thought the benefits outweigh the risk.

After that a number of studies were conducted in order to explain the privacy paradox in the sectors such as commerce, social media and healthcare.

2.3 Trust on institution

2.3.1 The concept of trust

In information privacy sector, trust is mostly defined as belief on media or institutions that gather information. Some researchers classified trust as trust on the internet website, internet media, and personal attitudes toward trust. Others classified trust as trust on the system, trust on the service provider, and trust on policy. The concept of trust has been widely used to explain the users’ behavior in terms of information privacy.

Chapter 3. Research Model

3.1 Research Model

The purpose of this research is to analyze online education users' privacy perception and their behavioral intention to disclose information as reaction. To investigate the gap between perception and behavioral intention, privacy calculus model was used. In other words, privacy calculus model was used as a tool to explain the decision-making process.

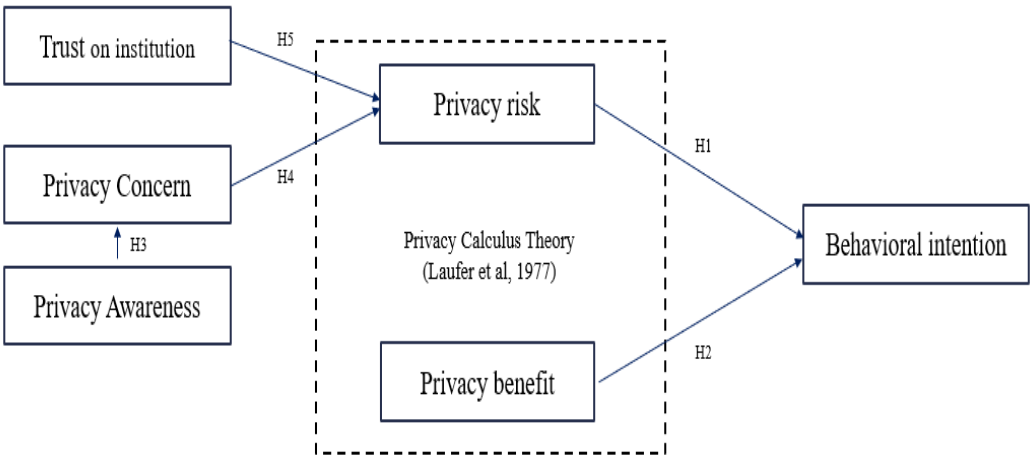

Figure 1. Research Model

3.2 Variables and Hypothesis

3.2.1 Operational definition of Variables

The variables used in this research are summarized in Table 3. The operational definitions of variables were made based on the ones verified in several previous studies.

Table 3. Operational definition of variables

Variable	Definition	Source
Behavioral intention	Intention to give information in order to use educational service	Malhotra (2004) Xu, Heng et al (2009)
Privacy Awareness	Extent to which an individual is informed about organizational privacy practices	Malhotra (2004) Hazari (2013)
Privacy concern	Uneasiness of releasing personal information	Smith et al.(1996) Dinev et al (2006)
Privacy risk	Degree to which users' belief that loss is possibly associated with the release of information	Malhotra et al (2000) Dinev et al (2013) Xu et al (2011)
Privacy benefit	Degree to which users' belief that beneficial outcome is associated with the release of information	Xu, Teo et al (2009) Xu et al (2011)
Trust on institution	Trust on the service provider (educational institution)	Dinev et al (2006) Krasnova et al (2010)

3.2.2 Hypotheses of this research

Based on literature review, variables, and research model, the hypotheses of this research are generated as follows.

H1: Users' privacy risk will negatively influence users' behavioral intention.

H2: Users' privacy benefit will positively influence users' behavioral intention.

H3: Users' privacy awareness will negatively influence users' privacy concern.

H4: Users' privacy concern will positively influence users' privacy risk.

H5: Users' trust in educational institutions will negatively influence users' privacy risk.

3.3 Research Methodology

In this study, empirical analysis was conducted to test the research model. For data collection, online survey was conducted. The survey was designed based on the constructs and items from previous studies. At least three items were included in each concept.

For the statistical analysis, SEM (Structural Equation Modeling) was applied. SPSS 25 and AMOS 25 programs were used for SEM and maximum likelihood

robust estimation. SEM analysis requires the reliability of the scale and the fitness of the measurement model (신건권, 2017). Therefore, each variable's measurement reliability and validity were tested by Cronbach's alpha, Composite Reliability (CR) and Average Variance Extracted coefficients (AVE). For the model fitness test, Confirmatory Factor Analysis (CFA) was employed. Last but not least, hypotheses was tested, using path coefficients and p-values.

Chapter 4. Empirical Analysis

4.1 Survey and data collection

For empirical analysis, this research gathered data by conducting an online survey. The survey was conducted for 5 business days via a professional survey firm in Korea. To avoid the concentration of particular gender and age groups, the proportion of gender and age groups was required to be equally distributed. As notified before, insincere or invalid responses were eliminated. As a result, total number of 319 responses were collected for the analysis. The demographics of this research sample is presented in Table 3.

Table 4. Demographics of research sample.

Demographic variables	Category	Frequency (Percent)
Gender	Male Female	160 (50.2%) 159 (49.8%)
Age	10s (14 ~ 19) 20s (20 ~ 29) 30s (30 ~ 39) 40s (40 ~ 49) 50s (50 ~ 59)	60 (18.8%) 65 (20.4%) 65 (20.4%) 64 (20.1%) 65 (20.4%)
Education Level	Below middle school High school Enrolled in college/university Graduated college/university Enrolled in graduate school Graduated graduate school	36 (11.3%) 53 (16.6%) 35 (11.0%) 164 (51.4%) 5 (1.6%) 26 (8.2%)

Monthly Income	Less than 1 million won 1 ~ 2 million won 2 ~ 3 million won 3 ~ 4 million won 4 ~ 5 million won 5 ~ 6 million won More than 6 million won	89 (27.9%) 33 (10.3%) 63 (19.7%) 53 (16.6%) 24 (7.5%) 31 (9.7%) 26 (8.2%)
Daily Internet Usage	Less than 1 hour 1 ~ 2 hours 2 ~ 3 hours 3 ~ 4 hours 4 ~ 5 hours More than 5 hours	27 (8.5%) 42 (13.2%) 65 (20.4%) 66 (20.7%) 42 (13.2%) 77 (24.1%)
Weekly Online Education Service Usage Time	Less than 1 hour 1 ~ 2 hours 2 ~ 3 hours 3 ~ 4 hours 4 ~ 5 hours	118 (37%) 91 (28.5%) 43 (13.5%) 28 (8.8%) 15 (4.7%)

4.2 Reliability and Validity Tests

Reliability test was performed to analyze how accurately and consistently the latent variable (or construct) has been measured. In this research, Cronbach's alpha was used to test the internal reliability of the latent variable. The measurement of latent variable is considered reliable when Cronbach's alpha's value is above 0.6. As the value becomes closer to 1, the measurement is considered more reliable. The constructs used in this research were found to be reliable as all the Cronbach's alpha's values were greater than 0.7.

Convergent validity was tested by measuring Composite Reliability (CR) and

Average Variance Extracted coefficients (AVE). The constructs were proved as valid as all the Composite Reliability values were above 0.7 and AVE values were greater than 0.5. The table 2 summarizes the result of reliability and validity test.

Table 5. Reliability and Validity test

Construct	Number of items	Cronbach's alpha	Composite Reliability	AVE
Privacy Awareness	3	0.826	0.853	0.660
Privacy Concern	4	0.918	0.936	0.786
Privacy Risk	4	0.799	0.848	0.588
Privacy Benefit	5	0.868	0.926	0.715
Trust on institution	4	0.834	0.897	0.687
Behavioral Intention	3	0.836	0.906	0.765

4.3 Confirmatory Factor Analysis

Next, Confirmatory Factor Analysis (CFA) was conducted to test the fitness of the model in this research. There are two types of factor analysis which are Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA is conducted when there is not any hypothetical or theoretical verification of construct in previous studies. However, as the constructs in this research is

verified in previous studies, only CFA was conducted.

The summary of model fit measures is presented in table 3. The model was proved to be fit in this research. Although the P-value was smaller than recommended value, other measurement values satisfied the recommended criteria. The value of χ^2/df was smaller than 3, GFI (Goodness of Fit Index) and AGFI (Adjusted Goodness of Fit Index) was greater than 0.9, 0.85 respectively. RMSEA (Root Mean Square Error of Approximation) was below 0.08 and NFI (Normed Fit Index) was greater than 0.9, proving the fitness of model.

Table 6. Model Fit Measures for Confirmatory Factor Analysis (CFA)

Fit measures	**Estimate**	**Recommended Values**	**Result**
χ2	262.017	-	-
df	109	-	-
χ^2/df	2.404	<3	Acceptable
P-value	0.000	>0.05	-
GFI	0.912	>0.9	Acceptable
AGFI	0.876	>0.85	Acceptable
RMSEA	0.066	<0.08	Acceptable
NFI	0.916	>0.9	Acceptable

4.4 Structural Equation Modeling and Hypotheses test

In this section, SEM analysis and hypotheses testing were conducted. The results are presented both on Table 7 and Figure 2. On Table 7, the estimates of regression weights and results of hypotheses testing is presented. The p-values were used to confirm the level of significance. Standardized estimates (β) were used as path coefficients. Based on this information, all the hypotheses except for hypothesis 5 were found to be accepted.

Hypothesis 1 is accepted as privacy risk (PR) gives negative effect on behavioral intention (BI) ($\beta = -0.215$, $p<0.001$). Hypothesis 2 is accepted as privacy benefit (PB) gives positive impact on behavioral intention (BI) ($\beta = 0.534$, $p<0.001$). Hypothesis 3 is accepted as privacy awareness (PA) gives negative effect on privacy concern (PC) ($\beta = -0.290$, $p<0.001$). Hypothesis 4 is also accepted as privacy concern (PC) gives positive effect on privacy risk (PR) ($\beta = 0.643$, $p<0.001$). Hypothesis 5 is rejected as the P-value is 0.772, proving that it is not statistically significant.

Table 7. Result of hypotheses testing by SEM analysis

Hypothesis & Path	Standardized Path coefficient	P-value	Hypothesis
H1: Privacy risk → Behavioral Intention	-0.215	***	Accepted

H2: Privacy benefit→ Behavioral Intention	0.534	***	Accepted
H3: Privacy awareness → Privacy concern	-0.290	***	Accepted
H4: Privacy concern → Privacy risk	0.643	***	Accepted
H5: Trust on institution → Privacy risk	0.531	NS	Rejected

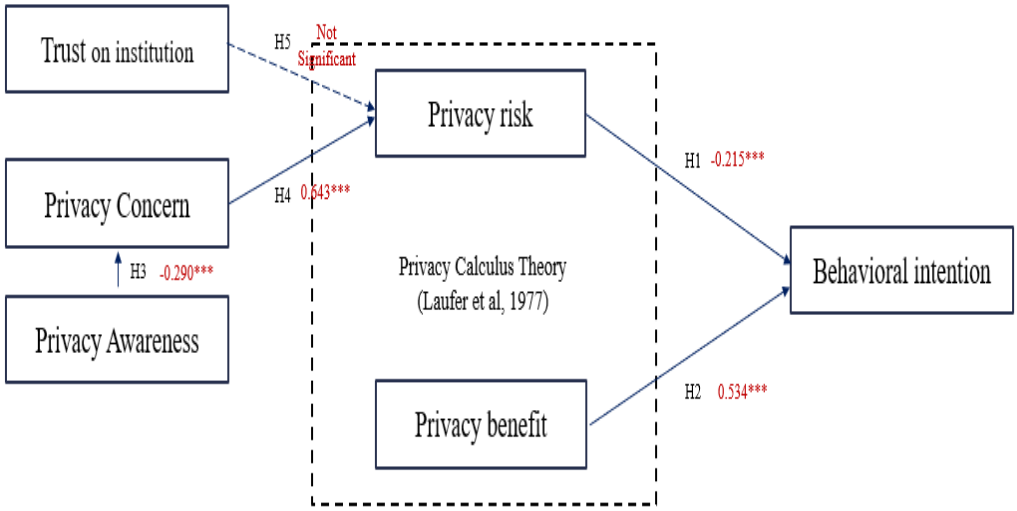

Figure 2. Path analysis result of research model

Chapter 5. Conclusion

5.1 Summary of research

The effort to provide education, anytime, anywhere, regardless of time and place have long continued. This form of education is called as ‘distance education’ and ‘online education’ is one type of distance education that has developed during the development of technology. And this dissertation examined online education users’ paradoxical behavior, where even though they are worried about disclosing their information, they decide to release the information. In order to explain this ‘privacy paradox’, the privacy calculus model was utilized, enabling the comparison between privacy risk and privacy benefit. Moreover, constructs such as privacy concern, privacy awareness and trust on institution were additionally examined for further investigation on the privacy calculus mechanism. The data for this research were gathered by online survey company, gathering 319 responds with almost equal proportion of each gender and age groups. The analysis results revealed that there are privacy paradox in online education sector with privacy calculus theory explaining why the respondents decided to do so. Except for the hypothesis on the relationship between trust and privacy risk, all of the hypotheses were accepted.

5.2 Implications

5.2.1 Academic Implications

The academic implication of this research is that it investigated users' privacy perception and their behavioral accordingly in the context of online education. Although collection and analysis of data in online education sector is becoming more common and widely conducted, only a few studies investigated on privacy issue in education sector. Therefore, this study scrutinized the privacy issues, especially privacy paradox, using privacy calculus theory.

5.2.2 Practical Implications

There are several practical implications of this study. First of all, this study revealed that educational service users have privacy concern when they disclose their data. Unlike the common belief that educational service users will be more comfortable in disclosing their data, they also had concern in giving their data to the service providers.

Second, despite their concern on privacy, the service users tend to disclose their data to use the service. By this, we were able to prove that there are paradoxical phenomenon in education sector also.

Last but not least, as the education sector is expected to gather more data with the help of technologies such as AI and learning analytics, the analysis on users' perception and behavior will help entrepreneurs and policy makers. By referring

to the education service users' perception and their behavior, it would be easier to design and develop better services.

5.3 Limitations and Future research

Although this research had academic and practical contributions, there are also several limitations of this study. First, although there are many types of online education services provided from universities to private companies, the analysis of this study was not able to differentiate the services. The survey respondents' reaction might have been different according to the characteristics of the educational institution. Therefore, the future study may improve the analysis by differentiating the characteristics of service providers. Another framework may be the key to analyze the difference of educational platforms.

Second, humans do not always make rational decisions. Privacy calculus theory is based on the belief that humans are rational. However, as several theories that tries to explain privacy paradox shows, humans can make biased decisions or decided to make easy decision based on heuristics. For further study, using another theory which is based on the belief that humans are not always rational may bring another theoretical, practical implications.

Bibliography

Badrul Khan. (2005). *Managing e-learning: Design, delivery, implementation, and evaluation*. IGI Global.

Brown, B. (2001). Studying the Internet Experience. *HP Laboratories Technical Report HPL, 49*.

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1). https://doi.org/10.1287/isre.1060.0080

Gilhooly, K. (2001). Making e-learning effective. *Computerworld*, *35*(29), 52–53.

Grubel, H. G. & Lloyd, P. J. (1975). *Intra-industry Trade*. London: Macmillan Press.

H. Jeff Simth, Tamara Dinev, & Heng Xu. (2011). Information Privacy Research: An Interdisciplinary Review. *Mis Quarterly*, *35*(4), 989–1015. https://www.jstor.org/stable/41409970

Henderson, J. M. & Quandt, R. E. (1987). Financing structure. *American Economic Review*, 39(1), 123-145.

Jones, K. M. L., Asher, A., Goben, A., Perry, M. R., Salo, D., Briney, K. A., & Robertshaw, M. B. (2020). "We're being tracked at all times": Student perspectives of their privacy in relation to learning analytics in higher education. *Journal of the Association for Information Science and Technology*, *71*(9), 1044–1059. https://doi.org/10.1002/asi.24358

Kenneth C. Laudon, & Jane P. Laudon. (2016). *MANAGEMENT INFORMATION*

SYSTEMS (14th ed.). Pearson.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers and Security*, *64*, 122–134. https://doi.org/10.1016/j.cose.2015.07.002

Li, Y. (2012). Theories in online information privacy research: A critical review and an integrated framework. *Decision Support Systems*, *54*(1), 471–481. https://doi.org/10.1016/j.dss.2012.06.010

Moore, M. G. ., & Greg Kearsley. (2011). *Distance education: A systems view of online learning*. Cengage Learning.

Nada Dabbagh, & Brenda Bannan-Ritland. (2005). *Online learning: Concepts, Strategies, and Application*. Prentice Hall.

Rita C. Richey. (2013). *Encyclopedia of Terminology for Educational Communications and Technology*. New York: Springer.

Rosenberg, M. J. . and R. F. (2002). *E-learning: Strategies for delivering knowledge in the digital age*.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., & et al. (2011). Open Learning Analytics: an integrated & modularized platform. *Open University Press*.

Szabo, M. (2002). Cmi theory and practice: Historical roots of learning managment systems. In G. H. and H. E. (pp. 929-936 n E-Learn: World Conference on E-Learning in Corporate (Ed.), *In E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 929–936).

Association for the Advancement of Computing in Education .

Watson, W. R., Lee, S., & Reigeluth, C. M. (2007). Learning Managemnt Systems: An Overview and Roadmap of the Systematic Application of Computers in Education. *Advances in Computer-Supported Learning*, 66–96.

Watson, William R., & Watson, S. L. (2007). An argument for clarity: What are learning management systems, what are they not, and what should they become? *TechTrends*, *51*(2), 28–34. https://doi.org/10.1007/s11528-007-0023-y

Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2009). The Role of Push-Pull Technology in Privacy Calculus: The Case of Location-Based Services. *Journal of Management Information Systems*, *26*(3), 135–174. https://doi.org/10.2753/MIS0742-1222260305

Young, J. R. (2019, December 4). New ownership for an LMS giant:Private equity firm to buy Instructure for $2 billion.EdSurge.Retrieved from https://www.edsurge.com/news/2019-12-04-new-ownership-for-an-lms-giant-private-equity-firm-to-buy-instructure-for-2-billion

Young, J. R. (2020, January 17). As instructure changes ownership,academics worry whether student data will be protected.EdSurge. Retrieved from https://www.edsurge.com/news/2020-01-17-as-instructure-changes-ownership-academics-worry-whether-student-data-will-be-protected

조은순, 엄명숙, & 김현진. (2012). *원격교육론*. 양서원.

한국교육공학회, 나일주, & 조은순. (2016). *교육 공학 탐구*. 박영사.

Moodle – Data privacy [Website]. (2021, Nov)

https://docs.moodle.org/400/en/Data_privacy

Appendix 1: Survey sheet

<온라인 교육 서비스 이용자의 개인정보 사용에 대한 인식 조사>

본 설문의 모든 내용은 무기명으로 처리되며, 통계적인 목적으로만 활용됩니다. 귀하의 귀중한 답변은 온라인 교육 서비스 이용자의 개인정보 사용에 대한 인식 분석 및 개인정보 이용 방안 연구에 소중히 사용될 것입니다.

- **온라인 교육이란?**

이러닝(e-learning)과 호환적으로 사용되기도 하는 이 개념은 전자 및 인터넷 기술을 이용한 교육을 지칭합니다. 지식과 수행을 향상시키기 위해 인터넷 기술로 다양한 유형과 범위의 학습활동과 자원을 전달(Rosenberg, 2001) 합니다.

- **프라이버시와 정보 프라이버시 개념**

프라이버시(privacy)란 다른 사람이나 기관 또는 국가의 어떠한 감시와 방해에서 자유롭기를 원하는 요구입니다(Kenneth C. Laudon, Jane P. Laudon, 2017). 여기서 더 나아가 정보 프라이버시란, 정보통신기술의 발달에 따라 기존의 프라이버시 개념에서 더 확장된 개념으로, 개인정보의 수집, 저장, 처리 및 배포 여부 및 방법 제어와 관련된 권리입니다(Spyros Kokolakis, 2017).

- **학습관리시스템과 시스템에서 수집되는 데이터의 종류**

학습관리시스템(Learning management system, LMS)이란 온라인 교육을 운영 및 관리하는 플랫폼으로, 교육의 개발, 전달, 평가, 관리 등 학습의 전반적인 과정을 관리합니다. 학습관리시스템에서 수집하는 데이터의 유형은 다음과 같습니다.

데이터 유형	유형별 수집 데이터
행정 데이터	신분, 신원 확인 데이터, 이미지 등
개인생활 데이터	생활양식, 가족상황 등
경제 및 금융 정보 데이터	소득, 재정 상황, 세금 상황
접속 데이터	IP 주소, 로그
교육 데이터	평가된 교과 과정, 시험 스크립트
학력 데이터	시험 성적, 평가, 수여 자격 등
위치 데이터	여행, GPS 데이터 등

Ⅰ. 설문 응답자 기본 정보 조사

1. 귀하의 성별은 어떻게 되십니까?
    - 남성
    - 여성

2. 귀하의 나이대를 체크해주세요
    - 10대 이상 20대 미만
    - 20대 이상 30대 미만
    - 30대 이상 40대 미만

- 40대 이상 50대 미만
- 50대 이상

3. 귀하의 최종학력은 어떻게 되십니까?
    - 중학교 졸업 이하
    - 고등학교 졸업
    - 대학교 재학
    - 대학교 (전문대) 졸업
    - 대학원 재학 (석, 박사)
    - 대학원 졸업

4. 현재 귀하의 월평균 소득은 얼마나 되십니까?
    - 100만원 미만
    - 100만원 이상 200만원 미만
    - 200만원 이상 300만원 미만
    - 300만원 이상 400만원 미만
    - 400만원 이상 500만원 미만
    - 500만원 이상 600만원 미만
    - 600만원 이상

5. 귀하의 1일 인터넷 사용 시간은 얼마나 되십니까?

- 1시간 이하
- 1시간 이상 2시간 미만
- 2시간 이상 3시간 미만
- 3시간 이상 4시간 미만
- 4시간 이상 5시간 미만
- 5시간 이상

6. 어떤 기관을 통해 온라인 교육 서비스를 이용하십니까? **(중복투표 가능)**
    - 대학 (온라인으로 전환된 수업)
    - 대학 (방송통신대학, 사이버대학 등 비대면 수업을 기본으로 하는 대학)
    - EBS, K-MOOC 등 국가 및 공공기관
    - 국내 사기업 (inflearn, NAVER edwith 등)
    - 해외 사기업 (Coursera, Udacity, edX 등)
    - 기타 : _____

7. 주로 어떤 내용의 온라인 교육 수업을 이용하십니까? **(중복투표 가능)**
    - 중등교육 (중고등학교 수업 내용)
    - 인문 분야 (언어, 역사, 철학 등)
    - 사회 분야 (경영, 경제, 심리학 등)
    - 공학 분야 (컴퓨터 공학, 코딩 등)
    - 공학 분야 (컴퓨터 공학, 코딩을 제외한 분야)

- 자연과학 분야 (수학, 통계학, 생물학 등)
- 의약 분야
- 교육 분야
- 예체능 분야

8. 1주일 간 온라인 교육 서비스 이용 시간은 얼마나 되십니까?
    - 1시간 이하
    - 1시간 이상 2시간 미만
    - 2시간 이상 3시간 미만
    - 3시간 이상 4시간 미만
    - 4시간 이상 5시간 미만
    - 5시간 이상

Ⅱ. 다음은 온라인 교육 서비스 이용 시 귀하의 개인정보 인식에 대한 설문입니다.

9. 온라인 교육 서비스 회사는 개인정보가 수집, 처리 및 사용되는 방식을 공개해야 한다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

10. 온라인 교육 서비스 회사/기관의 올바른 개인 정보 보호 정책은 명확하고 눈에 띄게 표시되어 있다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

11. 내 개인 정보가 어떻게 사용될지에 대해 나는 충분히 인지하고 있다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

Ⅲ. 다음은 온라인 교육 서비스 이용 시 귀하의 개인정보 제공 염려에 대한 설문입니다.

12. 나는 인터넷에서 제출한 정보가 오용될까 걱정된다.

- 매우 그렇다
- 그렇다
- 보통이다

- 그렇지 않다
- 매우 그렇지 않다

13. 나는 내 개인정보가 예상하지 못했던 방법으로 사용될 수 있다는 점이 걱정된다
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

14. 나는 다른 사람이 인터넷에서 내 사적인 정보에 접근할 수 있을까 걱정된다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

15. 나는 다른 사람들이 내 정보로 무엇을 할지 모르기 때문에 인터넷에 정보를 제출하는 것에 대해 걱정한다.
    - 매우 그렇다
    - 그렇다

- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

Ⅳ. 다음은 온라인 교육 서비스 이용 시 귀하가 생각하는 개인정보 제공으로 인해 발생할 수 있는 위험 사항에 대한 설문입니다.

16. 일반적으로 웹 사이트에 개인정보를 올리는 것은 위험을 수반한다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

17. 온라인 교육 서비스에 가입하고 이를 이용하면 개인 정보가 나도 모르게 사용될 수 있다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

18. 나의 정보가 온라인 교육 서비스 시스템에서 남용될 수 있다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

19. 온라인 교육 서비스를 이용하면 개인 정보의 사용에 따른 결과를 통제할 수 없게 될 가능성이 높아진다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

Ⅴ. 다음은 온라인 교육 서비스 이용 시 귀하가 생각하는 개인정보 제공으로 인해 발생할 수 있는 이익 사항에 대한 설문입니다.

20. 온라인 교육 서비스를 통해 내 학습 상태에 맞는 맞춤형 서비스를 제공받을 수 있다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

21. 온라인 교육 서비스는 내 학습 성향이나 개인적인 관심사에 맞춘 좀 더 관련 있는 교육 정보를 제공할 수 있다.

    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

22. 온라인 교육 서비스는 내가 좋아할 만한 교육 정보나 서비스를 제공할 수 있다.

    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

23. 온라인 교육 서비스를 통해 필요할 때마다 최신 교육 정보/서비스를 받을 수 있다.

    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다

- 매우 그렇지 않다

24. 온라인 교육 서비스를 사용하면 원하는 곳에서 관련 교육 정보/서비스에 액세스할 수 있다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

VI. 다음은 온라인 교육 서비스를 제공하는 회사/기관에 대한 인식과 관련된 질문입니다.

25. 온라인 교육 서비스 회사/기관은 신뢰할 만하다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

26. 온라인 교육 서비스 회사/기관은 개인 정보의 수집 및 사용 방침을 명확하게 공개한다.
    - 매우 그렇다
    - 그렇다

- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

27. 온라인 교육 환경은 정보 공개 및 교류하기에 안전한 공간이다..

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

28. 온라인 교육 서비스 회사/기관은 사용자들의 개인정보를 안전하게 보호한다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

Ⅶ. 다음은 개인정보와 관련된 귀하의 행동 경향 및 경험 그리고 학습과 관련된 질문입니다.

29. 온라인 교육 서비스를 이용하기 위해 개인정보 요구 시 나는 기꺼이 (willing) 정보를 제공할 것이다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

30. 온라인 교육 서비스를 이용하기 위해 개인정보 요구 시 나는 대체로 (likely) 정보를 제공할 것이다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

31. 온라인 교육 서비스를 이용하기 위해 개인정보 요구 시 나는 아마도 (probably) 정보를 제공할 것이다.
    - 매우 그렇다
    - 그렇다
    - 보통이다
    - 그렇지 않다
    - 매우 그렇지 않다

32. 개인정보 유출 경험이 있으십니까?

- 전혀 없다
- 직접 경험해봤다
- 간접적으로 경험해봤다 (가족 혹은 가까운 지인)
- 직/간접 모두 경험해봤다

33. 개인정보 유출/남용 사례에 대해 들어본 적이 있으십니까? **(중복 체크 허용)**

- 뉴스 검색을 통해 사례를 접했다
- SNS (카카오톡, 인스타그램 등)을 통해 들었다
- 대중매체 (TV 등)을 통해 접했다
- 다른 사람들의 개인정보 유출 사고에 관심 없다

34. 학습 계획을 세울 때 타인의 조언을 고려한다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

35. 교육 서비스를 제공하는 회사/기관의 정보 제공 동의 옵션에 대부분 동의하는 편이다.

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다

36. 나는 많은 사람들이 있는 공간보다 독립된 공간에서 혼자 교육을 받는 것을 선호한다.

1) 매우 그렇다
2) 그렇다
3) 보통이다
4) 그렇지 않다
5) 매우 그렇지 않다

Abstract (Korean)

온라인 교육은 원격 교육의 한 종류로 언제 어디서나 교육이 가능하게 한다. 그런 와중에 디지털 기술과 인공 지능, 학습 분석 등 보다 발전된 기술의 발달로 교육 분야의 데이터 수집 및 분석이 증가하고 있습니다. 그러나 교육 분야에서 프라이버시 문제를 조사한 연구는 많지 않아, 본 연구에서는 교육 현장의 프라이버시 문제를 분석하고자 하였다. 그 중에서도 특히 교육 서비스 이용자의 프라이버시에 대한 인식과 그에 따른 행동을 조사하였다.

본 연구에서 사용된 주요 이론은 프라이버시 역설과 프라이버시 계산이론이다. 프라이버시 역설은 개인이 자신의 정보를 공개하는 것에 대해 우려하면서도 자신의 정보를 공개하기로 결정하는 상황을 설명한다. 프라이버시 계산이론은 이러한 역설적 현상을 설명하는 이론 중 하나로, 개인이 무엇을 해야 할지 결정하기 전에 정보를 공개하는 것에 대한 이익과 위험을 경제적, 합리적으로 비교하는 이론이다. 분석을 위해 설문조사와 구조방정식 모형(SEM)을 실시하였으며, 가설 검정 결과 교육 부문에서도 프라이버시 염려가 있으며, 프라이버시 역설이 존재하는 것으로 나타났으며, 이중에서 프라이버시 계산이론은 개인이 자신의 정보를 공개하기로 결정하는 이유를 설명한다.

주요어 : 온라인 교육, 프라이버시 염려, 프라이버시 위험, 프라이버시 이익, 프라이버시 계산이론
학 번：2019-26284