

클라우드 컴퓨팅 서비스의 정보보호 실효성 증진을 위한 정보보호 정책의 언어적 특성 분석

정은한¹, 김경일^{2*}

¹(주)시스웨어 책임연구원, ²한국교통대학교 융합경영전공 교수

An Analysis of Linguistic Characteristics of Information Protection Policies to Improve the Effectiveness of Information Protection in Cloud Computing Services

Eun-Han Jeong¹, Kyung-Ihl Kim^{2*}

¹Senior Consultant, Sysware Co., Ltd.

²Professor, Division of Convergence Management, Korea National University of Transportation

요약 클라우드 서비스 제공자가 어떠한 정보보호정책을 소비자에게 제시하고 있는지 이용자는 잘 모르고 있는 것이 현실이다. 클라우드 서비스 공급자가 제공하는 정보보호 정책의 내용과 언어적 특성을 분석하여 정보보호의 실효성을 향상시킬 수 있는 방안을 찾고자 함이 본 연구의 목적이다. 연구 목적 달성을 위하여 클라우드 서비스를 제공하는 47개사의 정보보호 정책의 내용을 조사하고 언어적 특성이 미치는 영향을 분석하여 클라우드 서비스의 효율성을 높일 수 있는 있는 방안을 마련하고자 한다. 연구결과를 통해 기술적 처리방법의 포괄적 표현 등으로 인한 낮은 가독성이 법적 분쟁의 소지는 물론 클라우드 서비스의 확산에 방해가 될 수 있는 요인으로 나타났다. 연구결과는 이용자들에게 제공하여야 할 사항들을 제언함으로써 정보보호의 실효성을 높일 수 있을 것이다.

주제어 : 클라우드 컴퓨팅, 개인정보, 비밀성, 언어적 특성, 회계정보

Abstract It is a reality that users do not know well what kind of information protection policy the cloud service provider presents to consumers. The purpose of this study is to find a way to improve the effectiveness of information protection by analyzing the content and linguistic characteristics of information protection policies provided by cloud service providers. In order to achieve the purpose of this study, we investigate the contents of information protection policies of 47 companies that provide cloud services and analyze the influence of linguistic characteristics to come up with a plan to increase the efficiency of cloud services. The research results showed that low readability due to comprehensive expression of technical processing methods, etc., could lead to legal disputes and to hinder the spread of cloud services. The research results can increase the effectiveness of information protection by suggesting items to be provided to users.ing, Privacy, confidentiality, linguistic characteristics, Accounting Information

Key Words : Clouding Computing, Privacy, Confidentiality, Linguistic characteristics, Accounting Information

*Corresponding Author : Kyung-Ihl Kim(kikim@ut.ac.kr)

1. 서론

KPMG를 비롯한 많은 회계서비스 공급자들이, 다양한 클라우드 기반 회계정보시스템을 기업규모에 무관하게 제공하는 클라우드 서비스가 늘어나면서 클라우드 환경의 복잡성 또한 증가하였다.[1]. 프라이버시와 기밀보호는 클라우드 사용자와 공급자 모두에게 가장 중요한 문제이며 이는 클라우드 컴퓨팅 사용 의도를 저해시키는 요인이기도 하다[2]. 프라이버시와 기밀보호의 문제가 충분히 제시되지 않는 한 클라우드 컴퓨팅은 실패로 끝날 수밖에 없기에 클라우드 설계 모든 단계에서 검토되어야 할 사항이 된다[3]. 미국에서는 문제 해결을 위해 미국공인회계사회가 SOC(service organization control) 보고서 형태로 가이드라인을 제시하기도 하였지만, 여전히 회계담당자들은 제삼자의 회계데이터 관리라는 문제로 인하여 클라우드 기반 회계시스템 도입을 주저하는 것으로 보고된 바 있다[4].

사용자는 계약 협상에 앞서, 공급자의 정보보호 정책 검토에 시간과 노력을 기울일 것이기에 정책에 대한 설명과 내용은 잠재고객을 안심시켜서 마음을 끌 수 있는 중요한 요인이 될 것이다[5]. 본 연구는 서비스를 제공하는 클라우드 공급자 정보보호 정책의 내용과 불확실성, 언어의 어조 및 가독성과 같은 언어적 특성을 분석하여 공급자의 정보보호 정책의 목적 적합성을 확보함에 기초를 제공하고자 한다.

연구목적을 달성하기 위하여 정보보호 관련 정책에 대하여 그 내용을 5가지 영역으로 구분하여 영역별로 사용된 단어의 수, 단어의 성향 등 언어적 특성과 내용에 대하여 분석하였다. 아울러 사용자들이 읽고 이해하는 수준이 어느 정도 되는지에 대한 가독성을 영역별로 조사하였다.

연구의 심도를 높일 수 있도록 언어적 특성과 가독성의 상관분석을 통하여 정보보호 정책의 내용 중에서 공급자, 사용자 등 이해관계자들이 보다 많은 관심을 두어야 할 사항들을 찾고자 하였다. 아울러 서비스유형별로 언어적 특성과 가독성 분석을 시행하여 공급자 측면에서 고려해야 할 사항들을 제안하고자 하였다.

2. 배경이론 및 관련연구 검토

회계정보시스템은 클라우드 기술에 큰 영향을 받아 방대한 경영정보를 생성하여 의사결정을 지원하고 종업원과 고객을 원격에서 지원하는 선례가 없는 효익을

제공한다[6]. 미국의 경우 4대 회계 법인이 선도적으로 클라우드 서비스를 확장하면서, 2017년 3월에는 Deloitte가 Day1 Solution 클라우드 서비스를 개시하였고, KPMG와 MS사는 공동으로 회계감사용 클라우드 기반 플랫폼을 2017년 7월에 발표하기도 하였다[7].

Rong 등[7]은 클라우드 컴퓨팅의 가장 중요한 문제는 보안과 기밀보호라고 주장하였으며 공급자 시스템에 저장되고 처리되는 개인 데이터에 관한 관심은 클라우드 컴퓨팅의 가장 큰 장애요인이라는 주장도 있다[2].

클라우드 컴퓨팅은 다수의 사용자가 보편의 물리적 매체를 공유하기 때문에 데이터 보호에서 복수의 특성을 해결해야 하는 문제에 봉착한다[8]. 따라서 해커는 클라우드 서비스에 대해 단 한 번의 공격으로 다수의 사용자 정보에 접속할 수 있는 빈번한 사례가 있는바 그러한 공격들이 클라우드 컴퓨팅 정보보호의 문제를 증폭시키는 것이다[9].

사용자의 관심은 클라우드 공급자가 정보보호를 어떻게 하는가에 있다[11]. 공급자는 사용자에게 무슨 정보를 수집해서 어떻게 저장할 것이며 제삼자에게 무엇을 제공할 것인가와 같은 개인정보를 어떻게 다룰 것인가에 대한 정보보호정책을 제시한다[11]. 정보보호실행이 공급사가 제시한 정보보호정책을 위반했다는 관련된 법 조항과 예시를 정보보호정책에 기초해서 소송을 제기한 사례가 많이 있다[12].

이전 연구들은 조직의 정보보호정책은 소비자 정보보호에 대하여 중요한 역할을 하게 됨으로써 개인정보를 기꺼이 제공할 수 있다고 제시하고 있다[13]. 성인들의 인터넷 사용에 대한 조사에서 정보보호 위험에 관심이 있는 소비자는 정보보호 공개를 더 많이 읽을 것으로 조사되었다[5]. 또한, 정보보호 정책의 내용은 정보보호문제에 중요한 영향을 미친다는 연구[14]와 정보보호 정책 내용의 용어와 가독성은 중요한 편차를 보인다는 연구[13]도 있다.

현존하는 연구는 정보보호 정책이 사용자 인식과 행동에 의미 있는 영향을 미친다는 점을 강조하지만, 제한된 정보보호 정책에 기초한 연구라는 점이 제약된다. 클라우드 컴퓨팅 환경에서 중요한 정보보호 문제와 더불어 공급자가 정보보호 문제를 어떻게 취급하고 있고 정보보호 정책이 유효하게 전달 및 의사소통되고 있는지에 대한 체계적인 연구가 필요하다.

이하에서는 정보보호 정책의 주요 항목으로 본 연구가 채택하는 내용, 가독성, 불확실한 용어, 언어의 어조 등은 정보보호 위험에 대한 사용자 인식에 중요한 역할을 할 것이라는 가설을 제시하고자 한다.

3. 연구범위와 연구방법

3.1 연구범위와 문제의 제기

클라우드 서비스 공급자의 정보보호 정책은 FTC(Federal Trade Commission)의 프라이버시 원칙에 대처하도록 되었기에 그 5가지 영역의 내용은 논리적인 구조로 성문화 과정을 갖는다.[15]. 공급자의 내용 영역 분석은 다음과 같은 연구제기를 제시할 수 있다.

- 1) 어떤 정보를 수집할 것인가?
- 2) 누구에게 정보를 공개할 것인가?
- 3) 데이터 안전과 무결성을 유지하기 위해서 이용 가능한 측정방법은 무엇인가?
- 4) 데이터 보존과 이전에 관련된 절차는 어떻게 되는가?
- 5) 데이터 불만에 대한 절차는 무엇인가?

의사소통에 있어서 언어는 중요한 요인이 된다. 언어적 특징이 메시지에 대한 독자의 인식에 영향을 준다는 연구가 있다[16]. 본 연구에서는 불확실한 언어, 어조 및 가독성과 같은 정보보호 정책의 언어적 특성에 초점을 둔다.

첫 번째는 정보보호 정책을 전달하는 언어의 불확실성이다. 많은 연구는 문장 내 형식, 표현, 영향 및 태도를 조사하기 위해 불확실한 단어에 대한 지표를 사용했다[17]. 두 번째는 어조에 대한 것으로 부정적이거나 긍정적인 어조로 원문을 범주화한 연구로 부정적인 보고서는 긍정적 어조로 쓰인 보고서보다 더 강한 반응을 불러온다는 증거들이 제시되었다[18]. 세 번째는 가독성으로 공개나 보고서와 같은 문서의 가독성은 그 내용에 대한 사용자의 판단과 의사결정에 영향을 준다는 연구들이 있다[19]. 언어적 특성과 관련하여 다음과 같은 연구주제를 제시한다.

- 6) 클라우드 서비스 정보보호 서비스 공급자의 정보 보호 정책의 불확실한 언어 수준은 어떠한가?
- 7) 공급자 정보보호 정책의 언어적 어조는 어떠한가?
- 8) 가독성은 어떠한가?

본 연구에서는 5가지 내용 영역과 3가지 언어적 특성과의 관계를 규명하고자 동시 발생 빈도를 조사하고자 한다. 이는 내용상의 특성이 어떤 언어적 특성과 연

관될 수 있을 것이라는 생각에서이다.

- 9) 클라우드 서비스 공급자는 정보보호 정책에 있어서 특별한 언어적 특성을 이용하여 내용을 기술하는 것을 선호하는 것인가?

3.2 연구방법

제기된 연구를 수행하기 위하여 클라우드 서비스 공급사의 선정은 2019년 12월 말 현재까지 한국 클라우드 산업협회에 등록 및 가입한 임원사 14개사와 회원사 112개사 중에서 개인서비스가 아닌 비즈니스 서비스를 제공하는 47개사로 제한하여 직접 방문 및 온라인을 통해 조사를 수행하였다. 질적 내용분석을 위해서는 ATLAS-ti 소프트웨어 패키지를 이용하여 내용 영역과 언어 특성에 대한 데이터를 코드화하여 분석하였다.

내용 분석의 타당성과 신뢰성을 확보할 수 있도록 Chun[20]의 제안을 Fig. 1과 같이 인용하였다.

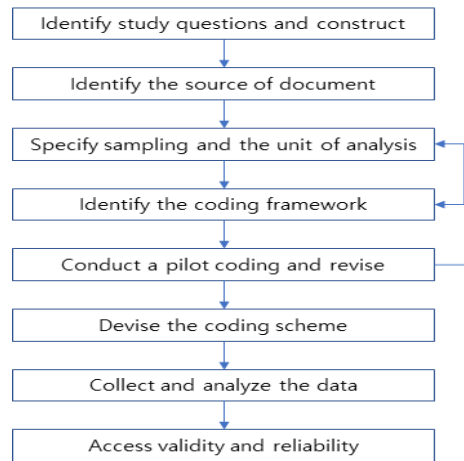


Fig. 1. Content Analysis Process

내용 영역을 분석하기 위해 Table 1에 제시한 바치럼 코딩 색인표에 따라 내용을 분류하였다. 언어특성을 살펴보고 불확실한 언어의 측정에는 ATLAS소프트웨어를 사용하였고 불확실성을 나타내는 단어에 대해서는 Loughran 등[21]의 연구를 인용하여 Loughran - McDonald Dictionaries로부터 단어목록을 사용하였으며 어조에 대한 측정은 Henry[18]의 연구를 인용하였다.

신뢰성검증은 Chua 등[13]의 연구를 인용하여 Flesch-Kincaid 신뢰성 테스트를 실시하였다.

Table 1. Coding Index for Format and Content Areas of Cloud Computing Service Providers' Privacy Policies

0. Privacy Policy Format 0-1 Bullet points 0-2 Q/A
1. Content Area 1: Data Collection(Q1:What information is collected?) 1-1 User information 1-2 Sever log or cookies 1-3 Service details 1-4 Mobile device 1-5 Details of the use of the information collected
2. Content Area 2: Data Disclosure(Q2: To whom is the information disclosure?) 2-1 Law and order 2-2 Business partners 2-3 Merger and acquisition
3. Content Area 3: Data Integrity(Q3: What measures are available to maintain data safety and integrity?) 3-1 General statement 3-2 Customers control their own data 3-3 List of procedures for customers to follow 3-4 Employee access restrictions 3-5 Employee confidentiality agreement and training 3-6 Location restrictions 3-7 Encryption or firewall 3-8 Testing of vulnerabilities 3-9 Notice of breach 3-10 Audited by auditors 3-11 Transparency report of law enforcement requests 3-12 Disclose to business partners with security and privacy measures
4. Content Area 4: Data Retention and Transfer(Q4: what are the procedures related to data retention and data transfer?) 4-1 Detailed data retention procedures 4-2 Data Transfer: ISMS 4-3 Data Transfer: APEC Framework
5. Content Area 5: Data Complaints(Q5: What are the procedures related to data complaints?) 5-1 Contact information provided 5-2 Investigate complaints and respond 5-3 Independent third party dispute channel available

4. 분석결과

조사대상 공급자 47개사 중에서 39개사는 비즈니스서비스 사용자만을 대상으로 하고 있고 8개사는 비즈니스와 개인 사용자 모두에게 서비스를 제공하는 것으로 파악되었다. 샘플 중 40개사는 서비스 유형이 한가지로 제한되었는데, 5개사는 IaaS, 8개사는 PaaS, 27개사는 SaaS를 제공하는 것으로 그리고 3개사는 IaaS와 PaaS를 4개사는 IaaS, PaaS, SaaS를 제공하는 것으로 조사되었다.

Table 2의 A항목은 정보보호 정책의 내용 영역별로 서비스 공급사가 계약내용에 포함하는 단어 수를 조사한 결과이다. Table 2의 B항목은 정보보호 정책의 내용을 영역별로 조사한 것으로 각 항목별로 분석된 결과를 설명하면, 조사대상 공급자들 중 89%는 사용자 이름, 비밀번호 같은 사용자 정보를 수집하고 있으며, 30%

는 사용자 서비스 사용 명세를, 21%는 사용자의 모바일 장비에 대한 정보를 수집하는 것으로 나타났다.

정보공개에 대한 조사 결과로, 81%가 제3자 비즈니스 파트너에게 공개한다고 하였는데 72%는 인수합병 시에 64%는 법원 등의 사법기관 요청이 있을 때에 공개한다고 하였다.

Table 2. Contents Analysis 1 in Cloud Service Providers' Privacy Policies

	A: Content Area Occurrence		B: Words Count	
	Count	Percent	Number of Words	Percent of Total Words
0. Privacy Policy Format				
0-1	47	100%		
0-2	13	28		
1. Data Collection				
1-1	42	89	4,983	10.20%
1-2	42	89	9,080	18.6
1-3	14	30	1,443	3.0
1-4	10	21	923	1.9
1-5	38	81	8,894	18.2
total			25,323	51.8
2. Data Disclosure				
2-1	30	64	1,978	4.0
2-2	38	81	4,483	9.2
2-3	34	72	1,786	3.7
total			8,247	16.8
3. Data Integrity				
3-1	32	68	1,306	2.7
3-2	10	21	618	1.2
3-3	9	19	494	1.0
3-4	11	23	493	1.0
3-5	4	9	131	0.3
3-6	7	15	181	0.4
3-7	16	34	520	1.1
3-8	2	4	13	0.03
3-9	3	6	190	0.4
3-10	1	2	54	0.1
3-11	2	4	99	0.2
3-12	13	28	644	1.3
total			4,743	9.73
4. Data Retention & Transfer				
4-1	26	55	2,495	5.1
4-2	32	68	4,427	9.1
4-3	2	4	80	0.2
total			7,002	14.4
5. Data Complaints				
5-1	42	89	2,253	4.6
5-2	10	21	367	0.8
5-3	18	38	944	1.9
total			3,564	7.3

데이터 안전과 무결성 유지를 위한 측정 결과를 분석하면 68%는 데이터 안전과 비밀을 보호 할 것이라는 일반적인 선언이며 34%는 데이터 암호화 및 방화벽 등의 구체적인 설명을 제공한다고 하였다. 권한 없는 접속에 대해서 23%가량이 종업원 접속 제한에 대하여 구체적인 언급이 있고 15%는 물리적 위치제한을 제시하였다. 9%만이 불법적인 고객 정보 유출로부터 근로자를 제한하는 교육과 비밀유지 동의서를 받는 것으로 나타났다. 사용자들은 클라우드 내 자신들의 데이터 통제권을 상실하는 것에 대한 염려가 있는데 21%만이 고객에게 권한이 있음을 천명하였고 19%는 고객이 그들의 데이터를 보호하는 절차에 대하여 제시하였다.

클라우드 플랫폼의 안전성을 증진하기 위해서 공급자는 실천을 검증할 수 있는 제3자 감사가 요구될 수 있지만 2%만이 이에 대해 언급할 수 있다. 시스템 안전 유지를 위해 4%는 예방조치활동과 주기적인 시스템 취약성 시험을 하는 것으로 나타났다. 6%는 고객에 대한 침해를 주지할 것이라고 했으며 4%는 사용자 데이터 접속을 위해 사법기관이 요청할 경우 사용자에게 투명하게 보고할 것이라고 하였다. 81%가 비즈니스 파트너에게 사용자 데이터를 공개할 것이라는 반면, 28%는 비즈니스 파트너에게 유사한 정보보호를 위한 측정을 하도록 요구할 것이라고 하였다.

데이터 유지와 이전 절차에 대한 조사결과, 55%는 고객이 서비스를 해지하고자 결정한다면 클라우드 내 데이터에 발생하는 바를 설명하는 구체적인 데이터 유지 절차를 가지고 있었다. 68%는 ISMS 표준을 이행하고 있으며 4%가량은 APEC 프레임을 준수한다고 명시하였다.

데이터 불만에 대한 조사에서 89%는 불만처리와 이해관계에 대하여 사용자에게 계약 정보를 목록화하였는데 38%는 공급자와 사용자 간의 분쟁을 해결하기 위하여 독립된 3자에 문제를 위탁할 수 있다고 명시하고 있다. 21% 정도만이 사용자 불만을 어떻게 조사하고 대응할 것인가에 대하여 명확히 하고 있다.

한편, 정책문서에 사용된 단어의 수와 비율에 대한 조사에서 51.8%의 단어가 데이터 수집에 대한 것이며 서버 로그와 쿠키에 대한 단어는 더 많이 이용되었다. 17%의 단어가 두 번째 영역인 데이터 공개에 대한 것으로 대부분의 단어들은 비즈니스 파트너에게 공개하는 것과 연관되어있다. 9.73%의 단어만이 데이터 무결

성에 대한 것으로 대부분 일반적인 선언에 불과한 것들이었다. 14%의 단어는 4번째 영역인 데이터 유지와 이전에 대한 것으로 사용되었으며 마지막으로 7.3%의 단어가 데이터 불만에 관련된 것으로 대부분 계약 내용에 해당되는 것이다.

Table 3은 불확실성을 내포하는 단어의 사용에 대한 조사결과로 불확실성을 내포하는 단어의 사용률은 극히 낮은 편으로 3.7%만이 사용되었다. 어조에 대한 특성을 살펴보았을 때는 긍정적인 어조도 부정적인 어조도 크게 발견되지는 않았는데 1%가량의 긍정적 어조의 단어와 0.2% 정도의 부정적 어조의 단어만이 발견되었다.

가독성에 대한 문제로서 이전의 연구들은 문서 대부분에서 60 이상의 Flesch Reading Ease, Level 7에서 8 이하의 Flesch-KInciad Grade를 갖는 문서가 잘 쓰이고 읽기에 쉽다는 견해를 보였다[15]. 본 연구에서는 분석결과 값이 평균적으로 FRE는 35.2, 등급 수준은 13.7로 나타나 조사 대상 공급자의 가독성은 매우 떨어져 읽고 이해하기에 난해한 것으로 나타났다.

Table 3. Linguistic Characteristics of Cloud Service Providers' Privacy Policies

Words Expressing Uncertainty and Linguistic Tone		
	Number of Words	Percent of Total Words
Total Words in the Policies	51,818	100.00%
Words Expressing Uncertainty	1,928	3.7
Linguistic tone words	600	1.2
Positive Tone Words	503	1
Negative Tone Words	97	0.2
Readability		
	Readability score	
Flesch Reading Ease	35.2	
Flesch-Kincaid Grade Level	13.7	

언어적 특성간의 상관관계를 분석한 결과는 Table 4에 제시하였는바, 불확실한 언어와 긍정적인 어조와의 관계는 유의미한 상관관계를 보였는데 이는 보다 긍정적인 단어를 사용하는 공급자는 불확실한 언어를 보다 많이 사용한다는 것을 의미한다. T-K등급수준과 불확실한 단어 간의 유의미한 상관관계가 뜻하는 바는 정보 보호 정책이 읽기에 어려운 내용으로 제공하는 공급자

는 불확실한 단어를 사용하는 경향이 있다는 것이다. F-K등급 수준과 FRE와의 '부'의 상관관계는 높은 등급이 높을수록 읽기에 어렵기 때문인 것으로 해석되어지는 당연한 결과로 보인다.

Table 4. Pearson Correlations among Linguistics in Cloud Service Providers' Policies

Variables	(1)	(2)	(3)	(4)	(5)
(1)Words Expressing Uncertainty	-				
(2)Negative Tone Words	0.16	-			
(3)Positive Tone Words	0.62*	0.12	-		
(4)Flesch Reading Ease	-0.28	0.12	-0.08	-	
(5)Flesch-Kincaid Grade Level	0.32*	-0.06	0.16	-0.73**	-

*,** Indicate correlation significance at the 0.05, 0.01 levels

5가지 영역별로 불확실을 내포하는 단어 사용에 대한 조사 결과는 다음 Table 5와 같다. 데이터 수집 영역에 있어서는 모바일 장비가 2.9%로, 데이터 공개 영역에서는 전체 평균 2.1%에 비하여 약간 상회하는 수준에서 법과 제도 부문이 가장 빈도가 높았으며 인수합병에서 가장 낮은 빈도를 보였다. 데이터 무결성 영역에서는 취약성 테스트 부문에서 가장 사용되었고, 암호화와 방화벽 분야에서 가장 적게 사용되었다. 데이터 유지와 이전 부문에 있어서는 데이터 유지 절차에서 가장 많이 그리고 APEC프레임에서 가장 적게 사용되었다. 마지막으로 데이터 불만 부문에서는 독립된 3자 분쟁 해결에 대한 내용에서 가장 많이 사용되었다.

Table 5. Word Expressing Uncertainty within Content Areas

	number	Percent of Total Words
1. Data Collection		
1-1	73	1.46%
1-2	134	1.48
1-3	26	1.8
1-4	23	2.49
1-5	144	1.62
total	418	1.65
2. Data Disclosure		
2-1	49	2.5
2-2	98	2.2
2-3	26	1.5
total	173	2.1
3. Data Integrity		
3-1	22	1.68
3-2	7	1.13
3-3	4	0.81
3-4	8	1.62

3-5	0	0
3-6	0	0
3-7	2	0.38
3-8	1	7.69
3-9	6	3.16
3-10	0	0
3-11	0	0
3-12	3	0.47
total	53	1.12
4. Data Retention & Transfer		
4-1	37	1.48
4-2	25	0.56
4-3	0	0
total	62	0.89
5. Data Complaints		
5-1	18	0.8
5-2	2	0.54
5-3	10	1.06
total	30	0.84

Table 6 에서는 언어의 어조에 관련된 결과를 제시 하는데 데이터 무결성 영역을 제외하고는 부정적 단어 보다는 긍정적 단어가 더 많이 발견되었다. 데이터수집 영역에서는 수집된 정보를 사용하는 내용에서 가장 많이 사용되어 공급자는 정보수집의 이유를 설명하고 정당화하기 위한 것으로 추론된다. 데이터 무결성 영역에서 부정적 단어가 상대적으로 많이 사용된 것은 데이터 무결성이 적절하게 보호되지 못했을 때 야기될 수 있는 부정적 결과를 설명함에 필요한 것이기 때문일 것으로 추론된다.

Table 6. Linguistics Tone within Content Areas

	positive tone		negative tone	
	number	percent	number	percent
1. Data Collection				
1-1	8	0.16%	4	8.00%
1-2	62	0.68	3	0.03
1-3	9	0.62	0	0
1-4	2	22	0	0
1-5	106	1.19	8	0.09
total	187	0.74	15	0.06
2. Data Disclosure				
2-1	11	0.56	4	0.2
2-2	19	0.42	1	0.02
2-3	1	0.06	1	0.06
total	31	0.38	6	0.07
3. Data Integrity				
3-1	2	0.15	3	0.23
3-2	1	0.16	2	0.32
3-3	0	0	2	0.4
3-4	0	0	1	0.2
3-5	1	0.76	0	0
3-6	1	0.55	0	0
3-7	1	0.19	1	0.19
3-8	0	0	0	0
3-9	0	0	0	0

3-10	3	5.56	0	0
3-11	0	0	0	0
3-12	0	0	0	0
total	9	0.19	9	0.19
4. Data Retention & Transfer				
4-1	2	0.08	2	0.08
4-2	7	0.16	1	0.02
4-3	0	0	0	0
total	9	0.13	3	0.04
5. Data Complaints				
5-1	2	0.09	0	0
5-2	0	0	0	0
5-3	1	0.1	1	0.1
total	3	0.08	1	0.03

Table 7에서는 가독성에 대한 결과를 제시하여 모든 영역에서 일반적으로 가독성이 낮은 것으로 나타났으며 데이터 수집 영역에서 다른 영역보다는 높은 가독성을 보였다. 데이터 수집영역에서 평균 FRE 37.8로 상대적으로 안정 값을 보이고 데이터 공개 영역에서도 평균 29.2로 큰 편차 없는 수준을 보였지만 데이터 무결성 영역에서는 큰 편차를 나타내어 가독성이 높이나 나타난 고객 준수 절차와 가독성이 가장 낮은 종업원 비밀보장 준수와 교육훈련에 대한 항목에서의 편차가 극심하게 나타났다. 데이터 유지와 이전 영역의 가독성 분석 결과는 APEC 프레임으로 데이터 이전하는 하위 항목에서 가장 가독성이 낮았다. 데이터 불만 영역에서는 3자 분쟁 조정 항목에서 가독성이 가장 낮게 그리고 불만 조사와 대응에 대한 항목에서 가독성이 가장 높은 것으로 조사되었다.

Table 7. Readability within Content Areas

	Readability	
	Flesch reading Ease	Flesch-Kincaid Grade Level
1. Data Collection		
1-1	34.5	14.60%
1-2	38.9	13.4
1-3	38.7	14.4
1-4	36.5	13.3
1-5	40.3	12.5
Mean	37.8	13.6
2. Data Disclosure		
2-1	29	16
2-2	3	15.2
2-3	28.8	16.8
Mean	29.2	16.0
3. Data Integrity		
3-1	10.5	17
3-2	38	13.3
3-3	45.2	114.4
3-4	19.2	17.4
3-5	0	18.4
3-6	30.5	12

3-7	32.8	13.5
3-8	19	13
3-9	30.6	14.8
3-10	20.9	14.8
3-11	24.4	15.5
3-12	23.7	17.8
Mean	24.6	14.9
4. Data Retention & Transfer		
4-1	41	12.7
4-2	22.7	16.5
4-3	3.5	17.6
Mean	22.4	15.6
5. Data Complaints		
5-1	26	14.7
5-2	33.1	13.2
5-3	19.9	17
Mean	26.3	15.0

추가적인 연구로 서비스 유형에 따라 내용과 언어적 특성의 차이가 존재하는지에 대하여 분석한 결과는 Table 8에 제시되었다. 사용된 단어의 수에 있어서는 비슷한 결과를 보였는데 전체적으로 IssS가 SaaS나 PaaS보다는 데이터 공개 영역을 제외하고 더 많은 단어를 사용하는 것으로 나타났으며 3가지 서비스 유형 간에 언어 특성상의 차이점은 발견하지 못하였다.

Table 8. Content Areas and Linguistic

Content Area	Mean	Word Count		
		IaaS (n=5)	PaaS (n=8)	SaaS (n=27)
Data Collection		704	510	583
Data Disclosure		106	177	208
Data Integrity		154	81	91
Data Retention & Transfer		157	94	185
Data Complaints		110	53	79
Total	1231.8	914.5	1129.2	
Linguistic Characteristics				
Uncertainty Language	41.2	43.4	41.3	
Positive Tone Words	10.6	9.1	10.6	
Negative Tone Words	2.2	1.5	1.6	
Flesch Readings Ease	28.3	31	36.7	
Flesch-Kincaid Grade Level	15.7	14.7	13.7	

5. 결론

본 연구에서는 클라우드 서비스 공급자의 정보보호 정책의 내용과 언어적 특성을 조사 분석하였다. 회계 관련 종사자, 잠재 및 현재의 고객들, 공급자와 정책입안자에게 중요한 연구결과를 제시할 수 있었다. 회계전문가들과 사용자들은 서비스공급자의 정보보호 정책에 향상된 방법으로 평가할 수 있어 사용자는 수집되지 않기를 원하는 데이터 수집에 대하여 정보보호 정책이 언급하는 것과 약 1/3 정도는 사용자의 서비스 이용에 대

하여 포함한 것 등을 예로 들 수 있다. 이러한 예는 사용자 프라이버시 침해에 해당하기에 사용자는 이러한 정보를 수집하지 않는 공급자를 찾을 수 있을 것이다.

서비스공급자에게도 좋은 제언이 될 수 있는 것은 가장 좋은 실천 방안을 제시하면서 개선의 필요성을 제시한 것이다. 무엇보다도 데이터 보안과 무결성은 클라우드 사용자에게 가장 중요한 것으로 데이터 안전과 무결성에 대한 일반적인 보장은 보편적으로 표현되고 있지만, 그 측정은 상세히 언급되지 않고 있다. 절반에 못 미치는 정책들은 상세한 보안 절차를 제시하였지만, 데이터 안전과 무결성 측정방법에 대해서는 전체 사용된 단어 중 9.7%만이 사용되었다. 더구나 데이터 안전과 무결성에 대한 절차의 기술은 다른 내용 영역보다도 낮은 가독성을 가진 것으로 나타났다. 공급자는 읽고 이해하기 쉬운 측정방법을 기술하여 고객을 유인하고 보장할 수 있어야 할 것이다.

정보보호 정책은 사용자가 서비스 이용 전에 반드시 동의해야만 하는 서비스라는 용어 정의가 가능하다. 공급자들이 자신들의 정보보호 정책을 위반하는 사례는 적지 않게 발생한다[14]. 그러나 정보보호 정책이 항상 법적인 구속력을 갖는 것이 아니기에 공급자의 침해로 인한 법적 대응이 능사가 될 수는 없다[18]. 포괄적인 용어와 불확실한 언어의 구사는 소송 위험에 대한 우려를 공급자가 안고 있음을 시사한다[10]. 그러나 소송에서의 낮은 성공률은 서비스 공급자가 낮은 가독성에 기인하는 비용 대비 소송 문제와의 균형을 맞출 필요성을 인식하여야 하며 불확실한 언어의 표현은 사용자 만족을 감소시키게 될 것이다.

사용자에게 보고되어지는 조사와 대응에 대한 정보는 극히 적어 89%가 불만 및 문제에 대한 계약 정보를 제공하고 있지만 21%만이 어떻게 조사하고 대응하여 보고할 것인가를 제시한다. 데이터 불만에 대해 어떻게 조사하고 대응할 것인가를 상세히 제공한다는 것은 고객의 만족을 증가시킬 수 있는 요인이 될 것이다.

불확실한 언어의 사용은 사용자 정보가 어떻게 공개되는가를 설명함에 보편적인 것으로 조사되었다. 사용자는 누가 자신의 데이터 접근하고자 하며 그 정보가 어디로 공개되어지는가에 관심이 있음에도 불확실한 용어로 인하여 이 문제는 더욱 증폭되어질 수 있다. 가독성이 낮다는 것은 전문 용어가 많고 공급자 간에도 그 용어가 달리 사용되는데 기인하는 것으로도 추론할

수 있다. 이 연구결과는 특별히 정책입안자들에게 의미가 있어 주요 용어의 표준화를 통한 정책 이해를 보다 용이하게 할 필요가 있다고 할 수 있다[20].

본 연구의 몇 가지 한계가 있다. 대부분의 계량적 연구들처럼 통계학적 중요성을 검증하지 않았다는 점이 우려될 수 있으나 클라우드 서비스 공급자의 정보보호 정책의 내용에서의 중요한 유형을 식별하고자 함에 그 의미를 더 많이 두고 있다. 또 다른 한계점은 47개의 표본에 대한 것으로 한국클라우드협회 회원사를 전수 조사하였다면 결과가 달리 도출되었을 수도 있었으나 연간 가입 회사와 도산 등의 이유로 탈퇴하는 회원사가 너무 많아 연구 한계를 안고 있을 수밖에 없었다.

마지막으로 정보보호 정책에 관한 기술은 실무에 적용함에 완벽하게 항상 정확하고 완전한 표현이 될 수는 없겠지만, 그 모순이 발생하는 한 이 연구는 계속되어야 할 필요가 있다.

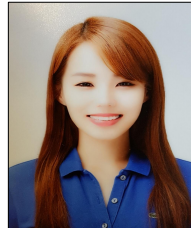
REFERENCES

- [1] C. Modi, D. Patel, B. Borisaniya A. Patel & M. Rajarajan. (2013). A Survey on security issues and solutions at different layers of Cloud computing. *The Journal of Supercomputing*, 62(2), 561-592. DOI : 10.1007/s11227-012-0831-5.
- [2] F. Shajzad. (2014). State of the art survey on cloud computing security. Challenges, approaches and solutions. *Procedia Computer Science*, 37, 357-362. DOI: 10.1016/j.procs.2014.08.053
- [3] S. Pearson. (2009). *Taking account of privacy when designing cloud computing services*. In 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing (pp. 44-52). IEEE. <http://www.hpl.hp.com/techreports/2009/HPL-2009-54.pdf>
- [4] J. J. Collins. (2015). *Online accounting systems: Accounting for cloud security*. <http://www.journalofaccountancy.com/issues/2015/sep/accounting-sytems-cloud-security.html>
- [5] G. R. Miline & M. J. Culnan. (2004). Strategiesofr reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive marketing*, 18(3), 15-29. DOI : 10.1002/dir.20009
- [6] P. Mell & T. Grance. (2011). *The NIST definition of cloud computing* <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistpecialpublication800-145.pdf>

- [7] C. Rong, S. T. Nguyen & M. G. Jaatun. (2013). Beyond lightning: A Survey on security challenges in cloud computing, *Computers & Electrical Engineering*, 39(1), 47-54.
DOI: 10.1016/j.compeleceng.2012.04.015
- [8] E. McCallister, T. Grance & K. A. Scarfone. (2010). *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII) (Vol. 800, No. 122)*. Diane Publishing, <https://csrc.nist.gov/publication/nistpubs/800-122/sp800-122.pdf>
- [9] A. K. VasanthaAzhagu & J. Gnanasekar. (2016). Cloud Computing Overview, Security Threats and Solutions-A Survey. *Proceedings of the International Conference on Informatics and Analytics*, Pondicherry, India, August 25-26.
- [10] H. Salcedo. (2014). *Google Drive, Dropbox, Box and iCloud reach the top 5 cloud storage security breaches list*, Hitachi <http://psg.hitachi-solutions.com/credeon/blog/google-drive-droopbox-box-and-icloud-reach-the-top-5-cloud-storage-security-breaches-list>
- [11] N. Steinfield. (2016). "I agree to the terms and conditions": (How do users read privacy policies online? An eye-tracking experiment, *Computers in Human Behavior*, 55, 992-1000.
DOI : 10.1016/j.chb.2015.09.038.
- [12] A. Seshagiri. (2013). Claims that Google violates Gmail user privacy, *The NewYork Times (October 1)*, <http://www.nytimes.com/interactive/2013/10/02/technology/google-email-case.html>
- [13] H. N. Chua, A. Herbland, S. F. Wong & Y. Chang. (2017). Compliance to personal data protection principles: A Study of how organizations frame privacy policy notices, *Telematics and Informatics*, 34(4), 157-170.
DOI : 10.1016/j.tele.2017.01.008.
- [14] K. W. Wu, S. Y. Huang, C. Yen & I. Popova. (2012). The effect of online privacy on consumer privacy concern and trust, *Computers in Human Behavior*, 28(3), 889-897.
DOI : 10.1016/j.chb.2011.12.008
- [15] S. Bradshaw, C. Millard & I. Walden. (2011). Contracts of clouds : Comparisonand analysis of the Terms and Conditions of cloud computing services, *International Journal of Law and Information Technology*, 19(3), 187-223.
DOI : 10.1093/ijlit/ear005
- [16] S. Goel & J. Gangolly. (2012). Beyond the numbers: Mining the annual reports for hidden cues indicative of financial statement fraud, *Intelligent Systems in Accounting, Finance & Management* 19(2), 75-89.
DOI : 10.1002/isaf.1326.
- [17] V. L. Rubin, E. D. Liddy & N. Kando. (2006). *Certainty identification in texts: Categorixation model and manual tagging results*. In *Computing Attitude and Affect in Text: Theory and Applications*, 61-76. Dordrecht, The Netherlands: Springer.
- [18] E. Henry. (2006). Market reaction to verval components of earnings press releases: Event study using a predictive algorithm, *Journal of Emerging Technologies in Accounting* 3(1), 1-19.
DOI : 10.2308/jeta.2006.3.1.1.
- [19] H. S. Asay, W. B. Elliott & K. Rennekamp. (2017). Disclosure readability and the sensitivity of invsetors' valuation judgements to outside information, *The Accounting Review* 92(4), 1-25.
DOI : 10.2308/accr-51570.
- [20] R. Chun. (2017). How virtuous global firms say they are: A content analysis of ethical values, *Journal of Business Ethics*, 155(1), 57-73.
DOI : 10.1007/s10551-017-3525-3.
- [21] T. Loughran & B. McDonald. (2014). Measuring readability in financial disclosures, *The Journal of Finance*, 69(4), 1643-1671.
DOI : 10.1111/jofi.12162.

정 은 한(Eun-Han Jeong)

[정회원]



- 2011년 2월 : 한남대학교 경영정보학과(경영학사)
- 2020년 2월 : 한국교통대학교 경영정보학과(경영학석사)
- 2020년 3월 ~ 현재 : 한국교통대학교 경영정보학과 박사과정 재학
- 2010년 12월 ~ 현재: (주) 시스웨어 재직
- 관심분야 : 중소기업정보화, 회계정보시스템
- E-Mail : eujung@sysware.co.kr

김 경 일(Kyung-Ihl Kim)

[종신회원]



- 1983년 2월 : 명지대학교 경영학과(경영학사)
- 1995년 2월 : 명지대학교 경영학과(경영학박사)
- 1993년 4월 ~ 현재 : (국립)한국교통대학교 융합경영학과 교수

- 관심분야 : 중소기업정보화, 회계정보시스템, IMS
- E-Mail : kikim@ut.ac.kr