Journal of Convergence for Information Technology
Vol. 12. No. 5, pp. 1-14, 2022
e-ISSN 2586-4440
DOI : https://doi.org/10.22156/CS4SMB.2022.12.05.001

지능정보사회의 안전한 인공지능 서비스 구현을 위한 개인정보 침해대응 및 위기관리 컴플라이언스 개발에 관한 연구

신영진
배재대학교 AI소프트웨어공학부 정보보안학 교수

A Study on Developing the Compliance for Infringement Response and Risk Management of Personal Information to Realize the Safe Artificial Intelligence Services in Artificial Intelligence Society

Young-Jin Shin
Professor, Division of AI Software Engineering-Information Security, PaiChai University

요 약 본 연구는 인공지능 서비스과정에서 개인정보를 포함한 데이터가 처리되고 있고, 그 과정에서 발생 가능한 개인정보 침해사고를 방지하기 위한 해결방안으로 개인정보 침해요인에 대응하는 위기관리 컴플라이언스를 마련하고자 한다. 이를 위해 먼저, 문헌조사 및 전문가 Delphi를 거쳐 처리과정을 범주화를 하였는데, 인공지능서비스 제공과정을 서비스기획·데이터 설계 및 수집과정, 데이터 전처리 및 정제과정, 알고리즘 개발 및 활용과정으로 구분하고, 3개 과정을 9단계의 개인정보처리단계로 다시 세분화하여 개인정보 침해요인을 구성하였다. 둘째, 조사한 개인정보 침해요인을 전문가 대상의 FGI, Delphi 등을 통해 선정하였다. 셋째, 각 개인정보 침해요인에 대한 심각도 및 발생가능성에 대해 전문가대상으로 설문조사하였으며, 94명의 응답결과에 대해 타당성 및 적정성을 검증하였다. 넷째, 인공지능 서비스에서의 개인정보 침해요인에 대한 적절한 위기관리 컴플라이언스를 제시하기 위해, 개인정보의 자산가치, 개인정보 침해요인, 개인정보침해사고 발생가능성을 활용하여 개인정보 침해 위험도 산정방식을 마련하였으며, 이를 통해 점수등급에 따라 위험정도에 따른 개인정보 침해사고 대응방안을 제시하였다.

주제어 : 인공지능 서비스, 개인정보보호, 개인정보 침해사고, 침해요인 심각도, 개인정보 위기관리 컴플라이언스, 델파이조사 심층면접조사

Abstract This study tried to suggest crisis management compliance to prevent personal information infringement accidents that may occur in the process because the data including personal information is being processed in the artificial intelligence (AI) service process. To this end, first, the AI service provision process is divided into 3 processes such as service planning/data design and collection process, data pre-processing and purification process, and algorithm development and utilization process. And 3 processes are subdivided into 9 stages following to personal information processing stages to infringe personal information. All processes were investigated with literature and experts' Delphi. Second, the investigated personal information infringement factors were selected through FGI, Delphi, etc. for experts. Third, a survey was conducted with experts on the severity and possibility of each personal information infringement factor, and the validity and adequacy of the 94 responses were verified. Fourth, to present appropriate risk management compliance for personal information infringement factors in AI services, a method for calculating the risk level of personal information infringement is prepared by utilizing the asset value of personal information, personal information infringement factors, and the possibility of infringement accidents. Through this, the countermeasures for personal information infringement incidents were suggested according to the scored risk level.

Key words : Artificial intelligence (AI) Service, Personal Information Protection, Personal Information Infringement Accident, Severity of Infringement Factors, Personal Information Crisis Management Compliance, Expert's Delphi, Focus Group Interview(FGI)

*This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2021S1A5A2A01069913)

*Corresponding Author : Young-Jin Shin(jinsyj@yahoo.com)

Received April 17, 2022 Revised May 06, 2022
Accepted May 20, 2022 Published May 28, 2022

1. 서론

지능정보사회에서 공공분야 및 민간분야에서 인공지능(Artificial intelligence, 이하 AI)서비스가 확산되고 있고, 교육, 의료, 교통, 민원 등 다양한 분야에서 활용되고 있다. 이를 위해 AI서비스 제공자는 지능정보기술을 활용하여 더 많은 서비스를 제공하기 위해 정보주체로부터 개인정보를 직접 수집할 뿐만 아니라, 수집된 데이터셋에서도 개인정보가 포함되기도 한다. 이로 인해 AI서비스 제공과정에서 개인정보 침해사고가 발생하고 있으나, 해킹, 악성코드 등 사이버공격에 대한 대응방안에 중점을 두고 있는 상황이다. 따라서, 본 연구에서는 AI서비스 제공과정을 구분하고, 그 과정을 개인정보처리단계로 세분화하여, 그 단계에서 발생할 수 있는 개인정보 침해요인을 해소할 수 있는 개인정보보호 방안을 마련하고자 한다.

물론, 현재 개인정보보호위원회(2021)가 AI서비스에서 개인정보의 처리흐름에 따라 단계별로 준수해야 할 AI 개인정보보호 자율점검표를 배포한 바 있다[1]. 그러나, 개인정보보호원칙에 중점을 두고 있기 때문에, AI서비스 제공과정에서 개인정보보호 준수사항으로 보기에는 한계가 있다. EU, 영국, 호주 등에서도 AI서비스에서의 개인정보 침해사고 대응을 위한 위험관리체계를 마련하고 있으나, AI서비스에 관한 책임원칙에 중심을 두고 있다. 따라서, 본 연구에서는 전문가대상으로 FGI/Delphi조사, 설문조사 등을 통해 AI서비스 제공과정에서 발생하는 개인정보 침해요인을 도출하고, 개인정보 침해위험도를 산정하여 침해위험정도에 따라 실행할 수 있는 위기관리 컴플라이언스를 제시하고자 한다.

2. 개인정보 침해사고 및 위기관리 프레임워크

2.1 개인정보 침해사고 사례

지능정보사회에서는 AI를 활용한 분야가 확대되고 있으며, 정보주체의 동의없이 AI서비스 제공과정에서 개인정보가 수집되는 경우가 발생하고 있다. 또한, AI서비스가 제공되는 과정에서 다양한 데이터가 결합하여 식별가능한 개인정보가 되고, 수집된 개인정보를 악용, 오·남용되기도 한다. 이에 대해 입법발전소(2018)는 개인정보가 AI기기에 의해 직접 수집되거나 개인사생활이 감시되기도 하며, 네트워크 연결센서를 통해 개인정보 및 사생활 침해가 발생할 수 있는 가능성을 제기하였다[2]. 즉, 지능정보기술이 탑재된 기기로부터 개인정보가 정보주체의 동의없이 수집되며, 명령어 오류로 인해 기기소유자의 명령없이 임의적으로 저장된 연락처에 이메일을 발송하는 사건도 발생한 바 있다[3]. 이외에도 적대적 스티커, 스피어싱, 사회공학적 공격자동화 등과 같은 개인정보 침해사고도 발생하였다[4]. 이처럼 AI서비스에서의 지속적으로 발생하는 개인정보 침해사고에 대응할 수 있는 개인정보보호방안이 필요하다.

2.2 개인정보 침해대응 위기관리체계

개인정보 침해대응 및 위기관리를 위한 점검체계에 관하여 국내외 사례를 검토하였다. 먼저, EU집행위원회는 AI 신뢰 확보를 위한 위험기반 AI시스템 규제체계(2021)를 제시하였다. 이는 AI시스템에서 발생할 수 있는 위험을 수용할 수 없는 위험(Unacceptable risk), 높은 위험(High-risk), 제한된 위험(Limited-risk), 최소한 위험(Minimal-risk)으로 구분하여, 위험수준별 위험관리방안을 제시하고 있다[6]. 영국의 UK Biometrics and Forensics Ethics Group(2019)은 실시간 얼굴인식기술로 인한 윤리문제를 해결하기 위해 법적 프레임워크를 개발하였으며, AI서비스에서의 개인정보 침해위협에 대해 발생규모, 피해정도 등을 고려한 위기관리체계로 활용하고 있다[7]. 호주의 산업과학에너지자원부는 AI시스템에 대한 위협요인의 발생가능성을 고려하여 위험평가 프레임워크(Assessment Framework for AI Systems: Risk)를 제시하였다[8]. 이는 개별 AI애플리케이션의 위험정도를 발생가능성(5단계), 위험심각성(5단계), 준수해야할 원칙(8가지), 위험정도(5단계)에 따라 대응활동을 구분하여 운영하고 있다. 이와 같은 호주의 AI 시스템에 대한 위험관리 프레임워크는 Table 1과 같이 정리하였다.

Table 1. Risk Assessment Framework for AI system in Australia

Div.		Details
Likelihood of risk (5)		·Rare, ·Unlikely, ·Possible, ·Likely, ·Almost certain
Consequence (5)		·Insignificant risk, ·Minor risk, ·Moderate risk, ·Major risk, ·Critical risk
Principle (8)		·Privacy Protection, ·Fairness, ·Physical harm, ·Contestability, ·Accountability ·Regulatory and legal compliance, ·Transparency and explainability, ·Number of people affected
Risk & Actions	Low	·Internal Monitoring ·Testing ·Review industry standards
	Moderate	·Internal Monitoring ·Consider how to lower risk ·Risk mitigation plan ·Internal review, ·Testing, ·Impact assessments · External review

(Continued)
Table 1. Risk Assessment Framework for AI system in Australia

Div.		Details
Risk & Actions	High	·Internal/External Monitoring, ·Consider how to lower risk, ·Risk mitigation plan ·Impact assessments, ·Internal and external review, ·Testing, ·Consultation with specialists, ·Detailed appeals/opt-out plan, ·Legal advice sought, ·Additional human resources to handle inquiries /appeals, ·Liaise with industry partners, government bodies on best practice
	Extreme	·Unacceptable risk

2.3 사전 연구

본 연구에서는 AI서비스 제공과정에서 발생할 수 있는 개인정보 침해사고를 방지하고자 개인정보 침해사고 대응 및 위기관리 컴플라이언스를 제시하여 개인정보보호를 강화하고자 한다. 이와 관련한 주제들의 연구들을 살펴보았는데, 김종락(2018)은 암호응용분야에 AI를 활용하여 사이버공격 이상행위탐지, 개인정보 오·남용 모니터링 등을 위한 보안기술을 강화하고자 하였다[9]. 최대선(2016)은 핀테크보안에서 활용할 수 있는 딥러닝기술을 검토하여, AI와 관련된 보안이슈, AI기술을 활용한 보안문제를 제기하였다[10]. 특히, 최대선(2017)은 주요국가에서 논의된 AI보안이슈에 관한 요인분석 및 대응사항을 Privacy by AI관점에서 제시하였다[11]. 홍은주 외(2019)는 신경망 학습단계, 예측단계 등에서 악의적 공격으로 인한 개인정보 노출을 해소하기 위해 신경망모델의 공격기법으로부터의 개인정보보호방법을 분석하였다[12]. 박소희 외(2017)는 AI기술을 악용하여, 데이터의 오작동, 학습데이터의 오염 및 탈취 등 사이버공격으로 인한 개인정보 이슈사항을 분석하여 개선하고자 하였다[13]. 박철희 외(2019)는 개인정보보호와 AI모델의 성능을 분석하여 개인정보보호를 위한 차분 프라이버시기술을 제안하였다[14]. L. Floridi(2018)는 AI기술의 악용으로부터 개인정보보호를 위한 20개 액션포인트를 평가하여, 책임원칙의 프레임워크, 감사 및 피해구제 메카니즘을 운영하고자 하였다[15].

그러나, 앞서 논의한 연구들은 AI로부터의 정보보안을 강조하고, 개인정보보호를 위한 안전한 관리체계의 중요성을 제기하지 못하였다. 물론, 신영진(2021. 3)은 인공지능서비스에서의 개인정보보호 이슈사항을 분석하고, 주요국가사례를 검토하여 법제적 기준 및 처리과정에서의 개인정보보호과제를 제시한 바 있다[16]. 특히, 신영진(2021. 6)은 주요국가에서 제시하는 AI준수사항에 관한 원칙을 비교하여, AI서비스에서의 개인정보보호를 위한 책임과 원칙을 강조하였다[17]. 이외에도 송기복 외(2020)는 AI서비스에서의 프로파일링을 위해 개인정보가 수집됨에 따라 개인정보 침해사고를 방지하기 위한 선제적 보호조치를 주장하였다[18]. 물론, 김종현 외(2016)[19], 한국인터넷진흥원(2020)[20] 등은 AI서비스 제공과정에서의 개인정보보호에 관한 컴플라이언스가 필요함을 강조하였다. 그러나, 개인정보처리단계에서 발생할 수 있는 개인정보 침해요인을 방지하기 위한 구체적인 방향을 제시하지 못하고 있다. 따라서, 본 연구에서는 사전연구를 참고하여 AI서비스 제공과정에서의 개인정보처리단계에서 발생할 수 있는 개인정보 침해요인을 도출하여 위기관리 컴플라이언스로 제안하고자 한다.

3. 연구의 분석 틀과 방법

3.1. 연구의 추진과정

본 연구는 AI서비스 제공과정에서의 개인정보 침해요인을 해소하기 위한 위기관리 컴플라이언스를 마련하고자 한다. 위기관리컴플라이언스는 개인정보 침해사고에 대응하는 매뉴얼, 안내서 등에서 보다 각 침해요인에 대처할 수 있는 세부적인 실행지침이라고 할 수 있다. 이를 위해서 먼저, 국내외 연구문헌, 국내외 보고서, 관련 법률 및 지침, 국제 표준 및 원칙, 인터넷 자료 등을 통해 국·내외 문헌조사를 하여, 개인정보 침해요인을 1차적으로 정리하였다. 둘째, 전문가대상의 FGI, Delphi분석을 하였다.[1] 특히, 개인정보 침해요인 도출과정에서 관련 전문가들을 대상으로 2021년 11월 9일, 12월 2~3일에는 FGI조사를 하여 개인정보 침해요인을 2차적으로 정리하였다. 또한, 관계전문가 6명을 대상으로 2021년 12월 18~20일, 12월 21~23일 온라인 Delphi조사를 하여, 개인정보 침해요인의 정제작업을 하였다.

더불어, 개인정보 침해요인에 대한 심각도, 발생가능성, 자산가치 등을 고려한 개인정보 침해요인의 위험도 산정방식은 개인정보 영향평가 수행안내서, 호주의 AI시스템에서의 위기관리컴플라이언스 등 기준을 정리하

1) · 심층면접조사(Focus Group interview: FGI): 소수관계자와의 집중화된 대화를 통한 면접방법으로 전문가집단토의, 집단면접 등으로 진행하는 조사방법
· 델파이조사(Delphi): 2~3회 전문가들의 의견을 듣고 피드백하여 예측하는 조사방법

였다. 이에 대해 전문가 Delphi분석을 통해 개인정보 침해대응 및 위험도를 산정의 적합성을 검증하고, AI서비스 제공과정에서의 개인정보 침해요인에 관한 심각도를 고려하여 위기관리 컴플라이언스를 제시하고자 하였다.

본 연구에서는 앞서 도출한 개인정보 침해요인의 중요도(심각도)에 관하여는 관계 실무자 및 전문가를 대상으로 설문조사(2022. 2. 3~2.15)를 하였다. 그 결과, 94명(학자, 실무자, 공무원 등)이 응답하였고, Table 2와 같다. 특히, 경력과 전공분야를 4개로 구분하였는데, 그 중, 11~20년 경력자가 34명, 21~30년 경력자가 23명 등 순으로 응답하였으며, 전공분야는 개인정보보호 및 정보보호분야 36명, 정책/법률 24명 등의 순으로 응답하였다.

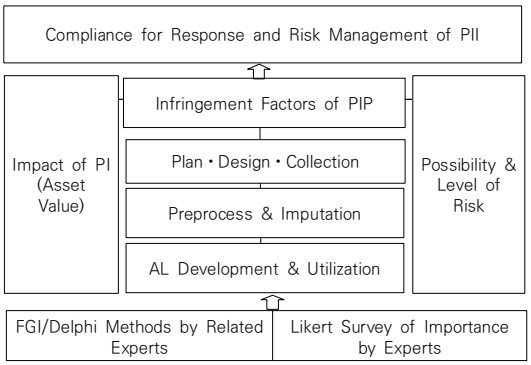
Table 2. Careers and majors of the experts

Year	1~10	11~20	21~30	31~40
People	18	34	23	6
Major	iCT	Privacy & Information Security	Policy/Law	Etc
People	17	36	24	5

3.2. 연구의 분석틀과 방법론

본 연구는 AI서비스 제공과정을 개인정보처리단계로 구분하고, 그 과정에서 발생할 수 있는 개인정보 침해요인을 도출하여, 그 과정에서의 개인정보 침해위험의 심각정도에 따라 대응할 수 있는 위기관리 컴플라이언스를 제시하고자 한다.

먼저, 본 연구의 핵심인 개인정보 침해위험요인은 관련 문헌검토, 전문가 FGI/Delphi분석 등을 통해 선정하였다. 이 과정에서 각 과정의 정의, 세부 구성요소를 추가 및 삭제, 카테고리의 범주화 등 각 침해요인의 적절성을 검토하였다. 둘째, 각 침해요인의 중요도은 위험심각성 및 발생가능성을 고려하였으며, 전문가 대상으로 5점척도 리커트 조사방식으로 설문조사하였고, 그 결과를 활용하여 각 침해요인을 SPSS를 활용하여 타당성 검토, 침해요인의 발생가능성 및 심각도에 관한 결과를 해석하였다. 셋째, AI서비스 제공과정에서 개인정보 침해위험도를 산정하여 침해위험도에 따른 대응방안을 제시하고자, 전문가 FGI/Delphi분석을 통해 산정방식, 산정요소, 대응사항 등을 검토하여 위기관리 컴플라이언스로 연계하였다.



* PI: Personal Information(hereafter PI)
PII: Personal Information Infringement(hereafterPII)
PIP: Personal Information Protection(hereafter PIP)

Fig. 1. Framework of the study

4. AI서비스 제공과정에서의 개인정보 침해요인 도출

4.1 개인정보 침해요인의 구성요소

AI서비스 제공과정에서 발생할 수 있는 개인정보 침해요인에 관하여는 개인정보보호위원회 외(2021), 한국정보화진흥원(2018), 신영진(2021.3) 등 관련 문헌들을 1차적으로 검토하였다. 이렇게 검토한 사항은 전문가대상의 FGI/Delphi분석을 통해 2차적으로 정리하여 선정하였다. 그 결과, 전체 AI서비스 제공과정은 서비스 기획·데이터설계 및 수집과정, 데이터 전처리 및 정제과정, 알고리즘(Algorithm, 이하 AL)개발 및 활용과정으로 구분하였으며, 각 과정을 세분화한 개인정보처리단계는 9개 단계이며, 각 처리단계에서의 침해요인은 58개 요인으로 구성하였다.

먼저, AI서비스기획·데이터 설계 및 수집과정(SDC)은 서비스 기획·데이터 설계단계(SP)에서 6개 및 데이터 수집단계(SC)에서 8개 개인정보 침해요인으로 구성하였다. 서비스 기획·데이터 설계단계는 AI서비스에 관한 개인정보영향평가(Privacy Impact Assessment: 이하 PIA) 및 정보주체의 동의를 받지 않고, 설계단계에서부터의 개인정보 침해위험을 제거하지 못하는 등 설계단계부터 준수해야 할 사항을 이행하지 않는 등 6개 침해요인을 선정하였다. 수집단계에서는 개인정보 수집시 최소한의 원칙을 준수하지 않고, 부적절한 접근방법, 정보주체에게 동의없이 수집하거나 고지하지 않고 수집하는 사항, 불법적 방법으로 AI 디바이스/제품에 대한 기기 및 사용자 인증없이 접속하거나, 송·수신시 개인정보 유·노출, 해킹을 통한 개인정보 변조, 접속인증 관리미흡으로 인한 제3자 무단접속

등에 관하여 8개 침해요인을 선정하였다. 이와 같이 AI서비스기획·데이터 설계 및 수집과정(SDC)에서 발생할 수 있는 개인정보 침해요인을 Table 3과 같이 정리하였다.둘째, 데이터 전처리 및 정제과정(DPI)은 전환·보유단계(8), 이용·제공단계(7), 위탁단계(2), 파기단계(4)로 구성하였다. 전환·보유단계(DT)는 AI디바이스/제품 등에 대한 보호조치를 수행하지 않고, 이로 인한 개인정보가 해킹되어 변조되거나, 가명정보 결합시 결합전문기관을 통하지 않고 처리하거나, 비인가된 AI디바이스/제품에 저장하여 개인정보가 유출되는 등 8개 침해요인을 구성하였다. 이용·제공단계(DU)는 AI 디바이스/제품에서 수집된 개인정보를 데이터 분석·가공을 위해 추출하거나, 본래 목적 외 동의없이 분석하거나, 추출된 개인정보와 타 기관 정보를 무단으로 결합하거나, 제3자에게 불법적으로 거래하는 등 7개 침해요인을 선정하였다. 위탁단계(DC)에서는 AI디바이스/제품에서 수집된 개인정보를 위·수탁시 정보주체에게 미고지하고 개인정보보호조치사항을 점검하지 않은 관리미흡을 침해요인(2개)으로 삼았다. 파기단계(DD)에서는 AI서비스 전처리 및 정제과정 이후 파기과정에서 제도적·기술적 보호조치가 미흡하거나, 지속적으로 개인정보를 보유하고 파기절차를 점검하지 않는 등 4개 침해요인을 선정하였다. 이와 관련하여 데이터 전처리 및 정제과정(DPI)에서의 침해요인은 Table 4와 같이 정리하였다.

Table 3. Components of PII Factors in Service Planning • Data design & Collection(SDC)

Stage	PII Factors	Ref.	Abbr.
Plan • Design (SP / 6)	It is not conducted Privacy impact Assessment for AI service	[1],[21]	SP1
	It is collected PI without the explicit consent of the AI device/application	[1],[21],[23]	SP2
	It is not removed the possibility of PII by reviewing in advance.	[16]	SP3
	It is not gathered opinions in the decision-making process among a series of processes for AI services	[22]	SP4
	It is limited the right to use AI service users	[22]	SP5
	It is not designed data sharing and response plans for emergencies when providing AI services.	[24]	SP6
Collection (SC / 8)	It is collected PI through inappropriate access to AI devices/products	[25]	SC1
	It is not notified of the standards for progressing PI for AI services (items not specified in PI, collection sources, indirect collection, etc.)	[25]	SC2
	It is not applied the standards regarding collection methods and protection measures when collecting PI	[25]	SC3
	It is collected information of Privacy monitoring without the consent of the information subject during providing AI service	[21],[25]	SC4
	It is accessible without device/user authentication for AI devices/products	[26],[27]	SC5
	It is transmitted/received PI(private information, transaction information, etc.) through connection with AI devices/products	[26]	SC6
	It is modified PI through hacking of access addresses with AI devices/products	[26]	SC7
	It is accessed by a third party without authorization due to insufficient management with AI devices/products	[28],[29]	SC8

Table 4. Components of PII Factors in Data Preprocess & Imputation(DPI)

Stage	Infringement Factors	Ref.	Abbr.
Transform • Possession (DT / 8)	It is leaked PI due to insufficient protection measures(encryption and access control for AI devices/products, management servers and application software, personal information processing systems, etc.).	[25],[30]	DT1
	It is processed pseudonymous information without going through a linkage information generation (combination key management) agency and a specialized binding agency, during combining pseudonymous information with PI collected from AI devices/products	[1],[21]	DT2
	It is modified PI by hacking the PI storage device collected from AI devices/products	[1],[21]	DT3
	It is not applied security techniques and protective measures(encryption of PI (sensitive information, etc.) collected from AI devices/products, protection measures for the database system where PI is stored, etc.)	[25],[30]	DT4
	It is leaked and exposed personally identifiable information collected from AI devices/products by AI systems/services	[25]	DT5
	It is leaked PI due to backed up to unauthorized devices (unauthorized and permitted devices) or stored in unauthorized AI devices and products	[24]	DT6
	It is not applied PI processing standards(de-identification, a combination of collected data, etc.) when converting to a dataset for AI learning according to data imputation and combination	[24]	DT7
	It is not applied quality labeling of AI applications applied to datasets including PI	[16],[24],[31]	DT8
Use·Offer (DU / 7)	It is extracted PI through analysis and processing of data collected from AI devices/products	[25]	DU1
	It is inappropriately analyzed personally identifiable information collected from AI devices/products for purposes other than the original purpose without the consent	[25]	DU2
	It is analyzed major databases of PI or sensitive information collected from AI devices/products without consent for purposes other than the original purpose	[8]	DU3

(Continued)

Table 4. Components of PII Factors in Data Preprocess & Imputation(DPI)

Stage	Infringement Factors	Ref.	Abbr.
Use·Offer (DU / 7)	It is not inspected PI processing status by combining personal information collected from AI devices/products with general information and de-identified information from other organizations	[8]	DU4
	It is illegally transacted such as transferring PI collected from AI devices/products to third parties	[25]	DU5
	It is provided PI without the consent of the information subject regarding the transmission of PI, provision to a third party, and use for other purposes, etc. when providing (selling) a third party for AI services	[1],[21],[32]	DU6
	It is not examed the basis and processing status of personal information transmission, provision to a third party, and use for other purposes, etc. when providing (selling) a third party for AI services	[20],[21]	DU7
Consignment (DC ./ 2)	It is not notified related matters, such as the content of work entrusted to the information subject when consigning or entrusting personal information collected from AI devices/products to a third party	[16],[20]	DC1
	It is insufficiently inspected and stored /managed PI protection measures (encryption of personal information, access control, etc.) when PI collected from AI devices/products are consigned to or entrusted to a third party	[16],[20]	DC2
Destruction (DD / 4)	It is insufficiently progressed institutional and technical protection measures in the process of destruction after preprocessing and inputting in AI services	[25]	DD1
	It is continuously retained PI even after the retention period in the pre-processing and inputting process in AI services	[25]	DD2
	It is not inspected the destruction procedure after the end of PI retention period in the pre-processing and refining process for AI service	[25]	DD3
	It is not examed the criteria for the destruction of PI and the retention period, even after achieving the processing purpose in the pre-processing and refining process for AI service	[25]	DD4

셋째, AL개발 및 활용과정(ADU)은 전환·보유단계(9), 이용·제공단계(9), 파기단계(5)로 구성된다. 먼저, 전환·보유단계(AT)에서는 AI 디바이스/제품 등에서 수집된 개인정보 보호조치, 개인정보 변조, 비인가된 제품 디바이스/제품의 개인정보 유출, AL 적용 범위 및 자동화된 의사결정에 대한 정보주체의 미동의, 데이터셋 구성에 대한 윤리성 미점검 등 9개 침해요인으로 구성하였다. 이용·제공단계(AU)는 AI서비스를 위한 AL개발을 위한 분석·가공을 위한 개인정보 추출, 수집된 개인식별정보의 부적절한 분석, 민감정보가 저장된 데이터베이스의 임의사용, AI를 위한 처리·분석·시각화과정에서의 개인정보 통제요인 미적용, 위험관련 정보 공개·공유, 개인정보처리과정 공개 등 개인정보처리지침 미반영 등에 관한 9개 침해요인으로 구성하였다. 파기단계(AD)는 파기과정에서 보호조치를 하지 않는 경우, 보유기간 종료 및 처리목적 달성 후에도 개인정보를 보유하는 경우, 데이터 파기시 파기준수사항을 미적용하는 등 5개 침해요인을 선정하였다. 이와 같이, AL개발 및 활용과정(ADU)에서의 침해요인 Table 5와 같이 정리하였다.

Table 5. Components of PII Factors in AL Development & Utilization(ADU)

Stage	Infringement Factors	Ref.	Abbr.
Transform • Possession (AT / 9)	It is occurred infringement (hacking, etc.) due to non-application of protection measures (encryption, stored data system protection, device, and user authentication) for PI collected from AI devices/products	[25],[33]	AT1
	It is modified PI by hacking PI storage devices collected from AI devices/products	[20],[24],[34]	AT2
	It is occurred infringement accidents (hacking, etc.) due to the non-application of security techniques in the process of developing and utilizing AI services	[29]	AT3
	It is exposed personally identifiable information to systems and services due to intentional or negligence in the development and use of AI services	[25]	AT4
	It is leaked PI due to backed up to unauthorized devices (unauthorized and permitted devices) or stored in unauthorized AI devices and products	[20],[24],[34]	AT5
	It is not applied de-identification measures such as personal identification and profiling when processing data through AI algorithm development and developed AI	[35]	AT6
	It is not consented the data subject to the scope of application of algorithms and automated decision-making according to automated data collection and storage including AI profiling	[36]	AT7
	It is occurred Infringement accidents due to the non-application of safe management protection measures such as encryption and access control	[25]	AT8
	It is not examined legal compliance and obligations when processing personal information for AI development (ex. unauthorized collection of personal information, big data collection method, legality of collection, de-identification of sensitive information, ethics in the composition of sensitive information dataset, etc.)	[20],[24],[34]	AT9

(Continued)

Table 5. Components of PII Factors in AL Development & Utilization(ADU)

Stage	Infringement Factors	Ref.	Abbr.
Use· • Offer (AU / 9)	It is extracted PI through analysis and processing for AI algorithm development	[16],[20]	AU1
	It is inappropriately analyzed personally identifiable information collected for AI algorithm development	[25]	AU2
	It is randomly used the main database of personal or sensitive information collected from AI devices/products	[22]	AU3
	It is analyzed habits and patterns without permission by combining PI with other general information to develop AI algorithms	[25]	AU4
	It is illegally transacted such as transferring PI to a third party during the AI algorithm development process	[25]	AU5
	It is leaked PI from web browsers, devices, and services that provide AI services	[22]	AU6
	It is possible to re-identify de-identified PI when processing big data through developed AI and providing it to a third party	[16],[20]	AU7
	It is not applied PI control factors in the process of processing, analysis, and visualization for AI	[21]	AU8
	It is not reflected in PI processing guidelines such as disclosure and sharing of risk-related information and disclosure of PI processing process	[21],[22]	AU9
Destruction (AD ./ 5)	It is insufficiently progressed institutional and technical protection measures in the process of destroying personal information in AI services	[25]	AD1
	It is continuously retained PI even after the personal information retention period in AI service	[25]	AD2
	It is deformed and retained PI due to insufficient control such as remote control of devices and services that provide AI services	[1],[21],[25]	AD3
	It is not destroyed PI even when the processing purpose is achieved, or the validity period ends	[1],[21]	AD4
	It is not applied the criteria for destruction (destruction plan after the achievement of purpose, irrecoverable destruction method, follow-up measures after destruction, etc.) after AI service is terminated	[1],[21]	AD5

Table 6. One-Sample Statistics and One-Sample Test of PII

Div	One-sample Statistics				One-sample Test					
	N	Mean	StDev	Se Mean	t	degree of freedom	significance probability	mean difference	95% Confidence Interval	
									lower	upper
PII	94	4.3756	0.38752	0.03997	108.222	93	〈0.001	4.32564	4.2463	4.4050
SDC	94	4.3393	0.42510	0.04385	97.827	93	〈0.001	4.28926	4.2022	4.3763
DPI	94	4.3655	0.43124	0.04448	97.023	93	〈0.001	4.31553	4.2272	4.4039
ADU	94	4.4221	0.43547	0.04492	97.341	93	〈0.001	4.37213	4.2829	4.4613

4.2 개인정보 침해요인 구성의 타당성 검증

본 연구에서는 AI서비스 제공과정에서의 개인정보 침해요인에 관한 중요도(심각도)를 관계 실무자 및 전문가를 대상으로 리커토 5점 척도를 활용한 설문조사를 하였다. 2022년 2월 3일부터 2월 15일까지 실시한 결과 94명이 응답하였으며, 그 결과를 SPSS프로그램, SmartPLS 등 통계프로그램을 활용하여 분석하였다. Table 6과 같이, SPSS를 이용하여 일표본 통계량 및 일표본 검정(검정값 =0.05)을 하였다. 즉, PII인 전체값에 대한 일표본 검정결과는 t=108.222, 유의확률 p=0.001(p〈0.05)로 유의미하였다. 이에 대해 비모수검정을 위한 일표본 카이제곱 검정(Chi-square test)을 하였는데, 검정통계량 x2=13.553로서, 유의확률 p=0.01(p.〈0.05)어서 유의미한 결과로 볼 수 있다. 이와 관련하여, AI서비스 제공과정의 개인정보 침해요인(전체합)과 각 서비스과정의 침해요인간의 상관성(C.I. Level=95.0)을 분석하였다. 그 결과, 서비스 기획 • 데이터 설계 및 수집과정(0.852), 데이터 전처리 및 정제과정(0.949), AL개발 및 활용과정(0.900)이었으며, Pearson상관관계를 분석한 결과, Table 7과 같이 높은 상관성을 확인할 수 있다.

Table 7. Pearson Correlation of composed process in PII

Div.	PII	SDC	DPI	ADU	Lower C.I	Upper C.I
PII	1.000	*0.852*	*0.949*	*0.900*	1.000	1.000
SDC	0.852	1.000	0.722	0.586	0.785	0.899
DPI	0.949	0.722	1.000	0.841	0.924	0.966
ADU	0.900	0.586	0.841	1.000	0.853	0.933

또한, SmartPLS를 이용하여 각 과정 및 단계의 경로분석을 하였는데, Fig. 2와 같다. 이에 대한 R

square=0.995인 개인정보 침해요인(0.995)에 관한 각 과정은 데이터 전처리 및 정제(0.412) 〉 AL 개발 및 활용(0.370) 〉 서비스 기획·데이터 설계 및 수집(0.310) 순으로 연관성이 있었다.

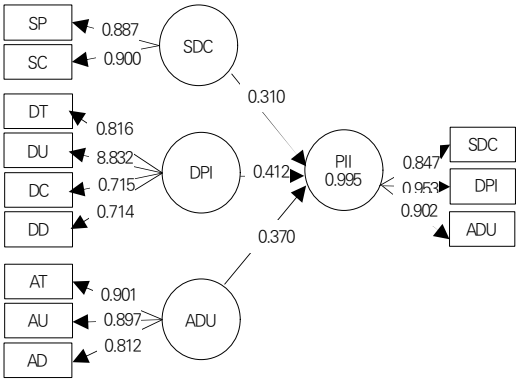


Fig. 2. Framework of the study

4.3 AI서비스에서의 개인정보 침해요인 심각도

4.3.1 개인정보 침해요인의 중요도

개인정보 침해요인의 중요도는 개인정보의 침해발생 가능성 및 개인정보보호 준수사항을 고려하여, 개인정보 침해사고의 발생가능성이 높은 경우, 그 피해규모가 크므로, 개인정보 침해요인의 피해 심각도로 해석할 수 있다. 이에 대해 관계 실무자 및 전문가를 설문조사를 하였다. 즉, AI서비스 제공과정에서의 개인정보 침해요인에 관한 중요도는 100%(분율)을 기준으로 82.26%였으며, 개인정보 침해요인이 발생하였을 경우 심각한 피해가 발생할 수 있다고 해석할 수 있다. 특히, 각 단계별 응답결과를 보면, AL 개발 및 활용과정 83.14% 〉 데이터 전처리 및 정제과정 82.07% 〉서비스 기획·데이터 설계 및 수집과정 81.58% 순으로 개인정보 침해요인의 심각성을 제기하였다. 이에 대한 응답결과의 타당성을 검증하였는데, Cronbach's Alpha=0.884로 신뢰도가 높게 나타났다.

첫째, 서비스기획・데이터 설계 및 수집과정에서의 개인정보 침해요인에 관한 설문조사 결과를 보면, 데이터수집단계(84.80%) 〉 서비스기획·데이터 설계단계(78.27%) 순으로 개인정보 침해요인의 심각성을 제기하였다. 특히, 서비스기획·데이터 설계단계(78.27%)에서는 AI디바이스/애플리케이션이 명시적인 정보주체 동의없이 개인정보 수집하는 것(84.00%)이 가장 위험하며, 데이터 수집단계(84.80%)에서는 AI서비스 제공 중 정보주체 동의없이 사생활을 모니터링한 정보를 수집하는 것(87.60%)이 가장 심각한 침해요인이 될 수 있다. 이외에 각 서비스기획・데이터 설계 및 수집과정을 구성하는 개인정보 침해요소의 중요도에 관한 설문결과는 Table 8과 같이 정리할 수 있다.

Table 8. Importance(severity) of PII in SDC

Div.	Importance (%)	Stage	%	Abbr.	%
SDC (2)	81.58	SP (6)	78.27	SP1	79.40
				SP2	84.00
				SP3	81.80
				SP4	73.60
				SP5	74.40
				SP6	76.40
		SC (8)	84.80	SC1	85.80
				SC2	81.40
				SC3	79.40
				SC4	87.60
				SC5	85.00
				SC6	87.20
				SC7	85.60
				SC8	86.20

둘째, 데이터 전처리 및 정제과정(82.07%)는 4개 개인정보 처리단계로 구성되는데, 각 단계별 개인정보 침해요인에 관한 중요도를 검토한 결과, 이용·제공(89.09%) 〉 전환·보유(87.35% 〉 파기(86.17%) 〉 위탁(86.49%) 순으로 개인정보 침해요인의 심각성을 제기하였다. Table 9는 데이터 전처리 및 정제과정에서 발생할 수 있는 개인정보 침해요인의 심각도를 조사한 결과이다.

Table 9. Importance(severity) of PII in DPI

Div.	Importance (%)	Stage	%	Abbr.	%
DPI (4)	82.07	DT (8)	82.17	DT1	86.20
				DT2	77.80
				DT3	84.40
				DT4	86.60
				DT5	86.40
				DT6	83.80
				DT7	80.20
				DT8	71.60
		DU (7)	83.74	DU1	81.40
				DU2	82.80
				DU3	83.80
				DU4	81.40
				DU5	88.60
				DU6	86.60
				DU7	81.60

(Continued)
Table 9. Importance(severity) of PII in DPI

Div.	Importance (%)	Stage	%	Abbr.	%
DPI (4)	82.07	DC (2)	83.74	DC1	80.00
				DC2	82.60
		DD (4)	81.00	DD1	82.60
				DD2	82.40
				DD3	79.00
				DD4	80.00

Table 9와 같이, 전환·보유단계에서는 AI디바이스/제품에서 수집된 개인정보 또는 민감정보에 대한 암호화, 개인정보 저장된 데이터베이스시스템의 보호조치 등 미이행(92.13%)이 가장 심각한 침해요인이다. 이용·제공단계에서는 AI디바이스/제품에서 수집된 개인정보 또는 민감정보의 주요 데이터베이스를 본래 목적 외로 동의없이 분석하는 경우(89.15%)가 가장 심각하다고 보았다. 위탁단계에서는 AI디바이스/제품에서 수집된 개인정보를 제3자 위·수탁시 처리되는 개인정보 보호조치(개인정보의 암호화, 접근통제 등)를 점검하지 않고, 보관·관리가 미흡한 경우(87.87%)가 개인정보 침해위험이 가장 높았다. 파기단계에서는 AI서비스를 위한 전처리 및 정제과정이후 파기과정에서 제도적·기술적 보호조치를 하지 않은 경우(87.87%)가 개인정보 침해위험이 높았다.

Table 10. Importance(severity) of PII in ADU

Div.	Importance (%)	Stage	%	Abbr.	%
ADU (3)	83.14	AT (9)	84.32	AT1	86.40
				AT2	86.40
				AT3	86.20
				AT4	8460
				AT5	83.20
				AT6	82.00
				AT7	81.00
				AT8	86.80
				AT9	82.20
		AU (9)	83.69	AU1	82.80
				AU2	82.60
				AU3	84.40
				AU4	83.20
				AU5	88.60
				AU6	85.20
				AU7	84.20
				AU8	81.20
				AU9	81.00
ADU (3)	83.14	AD (5)	81.40	AD1	81.20
				AD2	82.00
				AD3	82.00
				AD4	81.40
				AD5	80.40

셋째, Table 10과 같이 각 AL개발 및 활용과정에서 발생할 수 있는 개인정보 침해요인의 심각도를 검토하였는데, 개인정보 처리단계는 전환·보유단계(89.69%) 〉 이용·제공단계(89.03%) 〉 파기단계(87.59%) 순으로 개인정보 침해위험이 높게 나타났다. 각 단계의 침해요인을 살펴보면, 전환·보유단계에서는 암호화, 접근통제 등 안전한 관리 및 보호조치를 하지 않은 경우(92.34%)가 가장 심각한 침해요인이 될 수 있다. 이용·제공단계에서는 AL개발과정에서 개인정보를 제3자에게 양도하는 등 불법적 거래가 발생하는 것(94.26%)이 가장 심각한 침해요인이었다. 파기단계에서는 AI서비스를 위한 개인정보 보유기간이 종료된 후에도 지속적으로 개인정보를 보유(87.23%)하거나 AI서비스를 위한 디바이스 및 서비스 원격제어 등의 통제가 미흡하여 개인정보가 변형 및 미파기(87.23%)되는 경우 심각한 침해사고로 연결될 수 있다.

4.3.2. 개인정보 침해사고 발생가능성

앞서 논의한 AI서비스 제공과정에서의 개인정보 침해요인을 활용하여 개인정보 침해위험도를 산정하고자 한다. 이와 관련하여 호주의 발생가능성(5단계) 및 위험심각성(5단계), 국정원의 사이버 위기경보령(5단계), 개인정보보호위원회의 개인정보 침해위험정도(5점) 등을 활용하여 개인정보 침해사고의 발생가능성을 점수화하였다. 이와 관련하여 일본, 호주 및 우리나라에서의 개인정보 침해위험 발생가능성 및 위험특성을 고려한 위험도을 검토하여, 5단계(낮음, 높지 않음, 중간, 높음, 매우 높음)로 구분하여 1~5점을 부여하였으며, 그 내용은 Table 11과 같다[8,38,39].

Table 11. Possibility to Occur and Risk Degree of PII

Risk	Possibility of Occurrence	Characteristics of Risk	Risk Degree
Low or Minimal	Very rare	• It is not used PI or sensitive information • It is rarely possible to happen an incident of PII • It happens accidents with a low risk of affecting AI services.	1
Low–Moderate	Rarity	• It is used a small amount of PI in AI services • It is rarely possible to happen an incident of· PII • It may afftect accidents in AI services	2
Moderate	Possible	• It is used a large amount of personal information in AI services • It is possible to happen an incident· of PII • It happens accidents that likely to have a certain impact on AI services	3
High	Maybe· possible	• It collects PI without the consent specified in the AI application • It is possible to may occur immediately an incident of PII. • It happens accidents that are likely to· significantly affect AI services	4
Extreme	Almost certain	• It is used as the main database of PI/sensitive· information in AI services. • It is possible that may occur immediately a serious PII accident. • It happens accidents that are likely to seriously· affect many AI services.	5

Table 12. Impact of Personal Information (Asset Value)

Rank	Explanation	Asset Value
1st grade	• Identifiable PI or sensitive PI • PI that processing is strictly restricted in accordance with related laws • PI directly available to the crime in the event of a leak	5
2nd grade	• PI that can clearly identify an individual when combined • PI that can be held legally liable in case of leakage	3
3rd grade	• PI that can provide additional information in combination with personal information • PI that can be used illegally in limited areas	1

4.3.3. 개인정보의 영향도

앞서 본 연구에서는 AI서비스 제공과정에서의 개인정보 침해요인에 대한 심각성, 발생가능성 등을 검토하였는데, 개인정보의 침해요인의 심각성과 침해가능성이 높다는 것은 개인정보 침해위험수준이 높다고 해석할 수 있다. 그렇다면, 개인정보 침해요인이 될 수 있는 개인정보의 자산가치는 개인정보의 영향도로 해석된다[37]. 개인정보보호위원회 외(2020)는 개인정보 영향평가 가이드라인을 통해서 개인정보의 영향도 및 중요도에 대한 자산가치를 3등급으로 구분하여 5점(1·3·5)기준으로 점수를 부여하고 있다[37]. 이를 활용하여 본 연구에서도 동일한 기준을 적용하는 것이 적합한 지에 대해서는 전문가 Delphi분석과정에서 의견을 수렴하였다. 그 결과, 동일한 기준을 적용하여 객관성을 확보하는 것이 적절하다고 보았으며, 본 연구에서도 Table 12와 같이 본 개인정보의 영향도에 대한 등급을 적용하였다.

4.3.4. 개인정보 침해위험도 산정방식

개인정보보호위원회(2020)는 PIA를 위한 개인정보 침해위험도를 산정하고 있는데, 개인정보 영향도(자산가치), 침해요인 발생가능성, 법적 준거성 등을 고려하여 각 요소를 활용하여 산출하고 있다[37].[2] 이 산정방식의 활용에 대해서 전문가 Delphi분석과정에서 의견을 수렴하였는데, 기존 산출방식을 활용하는데 의견을 모았다. PIA의 개인정보 침해위험도를 산정하는 공식은 침해요인의 발생가능성과 법적 준거성을 곱하고 다시 2를 곱하는 방식이다. 이에 대해 AI서비스 제공과정에서의 개인정보 침해요인은 법적 준거성뿐만 아니라, 사전연구를 통해 개인정보처리를 위한 보호조치도 포괄하고 있기 때문에 법적 준거성을 대체하여 개인정보 침해요인의 심각도를 측정하였던 AI서비스 제공과정에서의 개인정보처리단계(9단계)별 각 1점을 부여하여 9점으로 배점하였다. 이에 따라 각 단계의 침해요인에 대해서는 1점을 기준으로 1/n(n=침해요인 발생수)를 적용하여 산정할 수 있다. 이렇게 개인정보 침해요인(N, 9점)과 발생가능성(B, 5점)을 곱하여 개인정보의 영향도(자산가치, A) 5점을 더한 값에 2를 곱한 점수(100점 기준)를 적용할 수 있다. 이러하 산정공식의 적

2) 개인정보 침해 위험도 = 개인정보 영향도(자산가치) + (침해요인 발생가능성 × 법적 준거성) × 2

합성에 대해서는 Delphi분석을 통해 검증하여 다음과 같이 도출하였다.

※ 개인정보 침해위험도 = {개인정보의 자산가치(5점) + (개인정보 침해요인(9점) × 발생가능성(5점))} × 2

$$= [A_k + (\frac{\sum_{i=1}^{N-M} k_i = k_1 + k_2 + \cdots + k_{N-M}}{N-M} + B_k] \times 2$$

- 개인정보 자산가치: A
- 개인정보 처리단계별 침해요인수: N
- 해당없는 침해요인수: M
- 개인정보 침해요인 발생가능성: B

4.4. 개인정보 침해대응 및 위기관리방안

이상과 같이 도출한 개인정보 침해위험도는 개인정보 침해위험으로부터 개인정보를 안전하게 관리하기 위한 대응방안을 마련하는 기준이 될 것이다. 앞서 호주의 경우 8가지 원칙에 관하여 5단계의 위험 발생가능성에 대한 위험수준을 구분하고, 위험정도별 보호조치를 권고사항으로 제시하고 있다[8].

이를 활용하여 Table 13은 개인정보 침해위험도를 점수화하여 5단계로 구분하고, 침해사고 대응 및 위기관리를 위한 컴플라이언스를 제시하였다. 물론, 법적 준수사항은 의무사항이지만, 그 밖의 침해위험에 대해서는 위험정도에 따라 단계별 대응조치가 개인정보보호를 강화할 수 있다. 따라서, 각 등급에 필요한 조치사항은 전문가 Delphi분석을 통해 Table 13과 같이 정리할 수 있다. Table 13은 AI서비스 제공과정에서 개인정보 침해요인이 발생하였을 때, 그 위험정도에 따라 적절한 대응을 할 수 있는 기준을 제시하고 있다. 즉, 개인정보 침해위험도가 심각한 경우(91~100점)에는 위험을 조절하거나 수용하기보다는 서비스자체를 종료해야 하는 상황이므로 대응방안을 제시하지 않았다. 그러나, 위험수준이 매우 낮은 최소한 위험이 발생하는 정도(0~20점)일 경우에는 개인정보 침해위험을 내부적으로 모니터링하고, 침해위험을 테스트하며, 개인정보보호지침을 검토하여 위험요인을 점검하여야 한다. 개인정보 침해위험수준이 낮은 경우(21~40점)에는 앞서 점검한 사항뿐만 아니라 개인정보 침해위험을 완화하는 계획을 수립하고, 개인정보 침해위험을 낮추기 위한 방안을 고려하여야 한다. 반면, 위험수준이 중간위험(41~75점)인 경우 앞서 점검한 조치 외에 AI서비스 기획·설계시 PIA를 수행하고, 외부전문가로부터 개인정보보호에 관한 사항을 검토를 받아 개선하여야 한다. 고위험(76~90점)인 경우는 중간위험일 때에 준수해야 할 사항뿐만 아니라, 개인정보보호를 위한 법률적 조언 및 외부전문가 컨설팅을 통해 개인정보 침해위험을 해소할 수 있다. 또한, 개인정보 침해위험에 대한 세부적인 이의제기·탈퇴계획을 수립하여야 한다. 더욱이, 개인정보 침해사고에 대한 문의 및 이의제기를 해결하기 위한 인력을 추가 배치하고 사고 대응을 위한 업계 파트너, 정부기관 등과 협력하여 문제를 해결해야 한다.

Table 13. Countermeasures by the Risk Level of PII

Div.	Risk	Detailed Actions
0~20	Low	· Internal monitoring of PII risk in AI service · Testing of PII in AI devices, media, service processing, etc. · Review of PIP guidelines for AI service
21~40	Low-Mioderate	· Internal monitoring of PII risk in AI service · Testing PII for AI devices, media, service processing, etc. · Review of PIP guidelines in AI service · Establishment of a risk mitigation plan for PII in AI services · Internal review of PIP in AI service
41~75	Moderate	· Application of PIA when designing AI services · Internal monitoring of PII risk in AI service · Tests of PII for AI devices, media, and service processing · Consideration of the ways to reduce the risk of PII in AI services · Establishment of risk mitigation plan of PII in AI services · Internal review of PIP in AI service · Review of PIP in AI service by external experts
76~90	High	· Application of PIA when designing AI services · Internal and external monitoring of PII risk in AI service · Consideration of the ways to reduce the risk of PII in AI services · Establishment of risk mitigation plan of PII in AI services · Tests of PII for AI devices, media, and service processing · Legal advice on PIP in AI services · Consulting on PIP in AI services by external experts · Detailed objection and withdrawal plan on PII risk in AI service · Addition of manpower to deal with inquiries and objections regarding PII risk and incidents in AI service · Liaise with industry partners and government agencies for best practices regarding PIP in AI services
91~100	Extreme	· Unacceptable risk

5. 결론

본 연구는 AI서비스의 기획단계부터 AL활용단계까지 전반적인 AI서비스 제공과정에서 발생할 수 있는 개인정보 침해요인을 개인정보처리단계별로 정리하였고, 각 침해요인으로 인한 개인정보 침해사고에 대응하기 위해 위험관리 컴플라이언스를 제시하였다. 즉, 본 연구에서는 전문가들의 의견을 수렴하여 체계화된 침해요인 및 대응기준을 마련하였고, 개인정보 침해요인의 중요도인 심각도를 관계 실무자 및 전문가를 대상으로 설문조사하였다. 각 AI서비스 제공과정 및 개인정보처리단계에서의 개인정보 침해요인에 대한 설문결과의 타당성 및 적절성을 SPSS를 활용하여 검증하였다. 또한, 개인정보 침해요인에 대응하기 위한 절차를 마련하기 위해, 개인정보 침해위험도를 산정하여 위기관리 컴플라이언스로 제시하고자 하였다. 즉, 개인정보의 영향도(5점)를 기준으로 개인정보 침해위험요인(9점)과 발생가능성(5점)과 연계하여 100점 기준으로 점수산정방식을 도출하였다. 이렇게 산정가능한 개인정보 침해위험도는 점수에 따라 5단계로 등급화하여 각 점수구간에서 개인정보 침해사고에 대응할 수 있도록 차별화된 위기관리 컴플라이언스를 제시하였다. 이를 활용하여 AI서비스 제공과정에서의 개인정보 침해사고에 대한 내외적인 대응체계, 대응기준, 대응시나리오 등을 마련할 수 있다.

앞으로 지금보다 더 많은 분야에서 AI서비스가 제공될 것이므로, AI서비스 제공과정에서의 개인정보보호는 무엇보다 중요한 과제가 될 것이다. 지금까지 개인정보 침해위험과 사이버위험을 구분하여 대응할 수 있는 관리기준이 모호하였으며, 개인정보보호의 범위규정도 모호하여 개인정보 침해사고에 대한 논란이 계속되고 있다. 이에 대해 본 연구에서는 AI서비스 제공과정에서 발생할 수 있는 개인정보 침해요인을 도출하고, 각 개인정보처리단계에서 발생할 수 있는 개인정보 침해요인을 해소하기 위한 위기관리 컴플라이언스를 제시하였다. 즉, 단순히 업무절차, 매뉴얼, 안내서 등으로의 가이드라인이 아니라 각 침해요인에 대해 적극적으로 대응할 수 있는 체계화된 위기관리 컴플라이언스를 제시하고자 하였다. 이러한 본 연구의 성과를 활용한다면, AI서비스가 신뢰받는 환경에서 구현될 것이며, AI서비스 제공자도 구체화된 개인정보보호조치를 실천할 수 있는 적극적인 대응체제를 마련할 수 있으리라 본다.

REFERENCES

[1] Personal Information Protection Committee. (2021. 5. 31). *Artificial Intelligence (AI) Personal Information Protection Self-Checklist*.

[2] The legislative power plant. (2018). *Overseas discussions on AI privacy issues*(Online). http://blog.naver.com/PostView.nhn?blogId=legislationpp&logNo=221408527453

[3] K. W. Kug. (2019. 3. 2). Application examples by Artificial intelligence technology and industry, *Weekly trend*, 15-27.

[4] R. Goosen, A. Rontojannis, S. Deutscher, J. Rogg, W. Bohmayr & D. Mkrtchian. (2018. 11. 13). *Artificial Intelligence Is a Threat to Cybersecurity. It's Also a Solution.* Technology & Digital BCG(Online). https://www.bcg.com/publications/2018/artificial-intelligence-threat-cybersecurity-solution.aspx

[5] National Information Society Agency. (2017). *Korea AI Company Status Survey Report,* 1-119.

[6] D. H. Kim. (2021. 5. 23). *AI guidelines' from around the world will compete.. A small step domestically.* News Tomato(Online). http://m.gobest.news.dreamwiz.com/NEWSAXmV7Is1a3AFLjqbk89k

[7] UK Biometrics and Forensics Ethics Group. (2019). Ethical issues arising from the police use of live facial recognition technology. *Interim report of the Biometrics and Forensics Ethics Group Facial Recognition Working Group.*

[8] Australian Government, Department of Industry, Science, Energy, and Resource. (2019). *Artificial Intelligence: Australia's Ethics Framework, A Discussion Paper,* 1-76.

[9] J. L. Kim. (2018). Relation between Artificial Intelligence and Information Security. *Communications of the Korean Institute of Information Scientists and Engineers. 36(2),* 14-17.

[10] D. S. Choi. (2016). Artificial Intelligence and Fintech Security. *Review of KIIC, 26(2),* 35-38.

[11] D. S. Choi. (2017). Artificial Intelligence and Security. *The Journal of The Korean Institute of Communication Sciences, 34(10),* 31-37.

[12] E. J. Hong, S. J. Lee, D. W. Hong & C. H. Seo. (2019). *Analysis of privacy issues and countermeasures in neural network learning, Journal of Digital Convergence. 17(7),* 285-292.

[13] S. H. Park & D. S. Choi. (2017). *Artificial*

Intelligence and Security Issues, Review of KIISC, 27(3), 27-32.

[14] C. H. Park & D. W. Hong. (2019). Differential Privacy Technology Resistant to the Model Inversion Attack in AI Environments. *Journal of the Korea Institute of Information Security & Cryptology, 29(6)*, 589-598.
DOI : 10.13089/JKIISC.2019.29.3.589

[15] L.Floridi et al. (2018). AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines. 28,* 689-707.
DOI : 10.1007/s11023-018-9482-5

[16] Y. J. Shin. (2021). The Improvement Plan for Personal Information Protection for Artificial Intelligence (AI) Service in South Korea. *Journal of Convergence for Information Technology. 11(3),* 20-33.
DOI : 10.22156/CS4SMB.2021.11.03.020

[17] Y. J. Shin. (2021). A Study on the Application of Responsibility and Principle for Personal Information Protection in AI Services. *Korea Criminal Intelligence Review, 7(1),* 45-74.

[18] G. B. Song & J. K. Lee. (2020). Discussions on the Commercialization of AI and the Protection of Personal Information - Focusing on Image Data and Profiling -. *Korean Security Journal,* 65, 453-476.

[19] J. H. Kim et al. (2016). *A study on the protection of personal information in the field of artificial intelligence and robotics*, Personal Information Protection Committee.

[20] Korea Internet & Security Agency. (2020). A Study on Improvement Plans for Personal Information Protection Policy for Fostering the Artificial Intelligence Industry. *Service Proposal Request Form.*

[21] Personal Information Protection Committee & Korea Internet & Security Agency. (2021). Artificial Intelligence and Personal Information Protection. *The 6th Pseudonym Information Expert Training Materials.*

[22] National Information Society Agency. (2018). Threats, and countermeasures against the exploitation of artificial intelligence. *NIA Special Report.*

[23] M. Comiter. (2019). Attacking Artificial Intelligence. *Harvard Kennedy School Belfer Center for Science and International Affairs, 2019-08.*

[24] Ministry of Science, Technology, and Information & National Information Society Agency. (2021). *A guide to building datasets for artificial intelligence learning.*

[25] Personal Information Protection Commission. (2015). *Analysis of personal information infringement factors in the Internet of Things era and investigation of actual cases.* Personal Information Protection Committee.

[26] H. J. Shim. (2018). *The Paradox of Artificial Intelligence (AI) and Privacy: Focusing on AI Voice Assistants.* KISDI Premium Report.

[27] B. G. Lee. (2021). *The final report on the development of AI dysfunction prevention technology by deceptive attacks.* Information and Communication Planning and Evaluation Institute, Ministry of Science and ICT

[28] D. H. Kim, S. W. Yoon & Y. P. Lee. (2013). Security for IoT Service. *The Journal of The Korean Institute of Communication Sciences. 30(8)*, 53-59

[29] J. H. Jeon. (2015). Analysis on the Security threat factors of the Internet of Things. *Convergence security journal. 15(7),* 47-53.

[30] Y. J. Shin. (2020). A Study on Developing and Applying Framework and Assessment Standard of Its Conformity of Personal Information Protection for IoT Service Subject. *Journal of Korean Association of Regional information Society, 23(2),* 83-117

[31] S. J. Sohn & S. J. Ahn. (2021). A Study on the Artificial Intelligence Ethical Principle Classification Model. *Proceeding of the Korean Association of Computer Education.* 25(2), 111-114.

[32] Korea Internet & Security Agency. (2021). Main contents of EU artificial intelligence (AI) regulation and protection of personal information. *Monthly* trend analysis of personal information protec*tion. 5,* 1-15.

[33] K. J. Choi. (2015). A Study for Developmental Change of Personal Information Protection Law System in the Age of Big Data and Internet of Things (IoT). *Chung_Ang Law Review, 17(4),* 7-50.

[34] Y. D. Kim and W. C. Jang. (2016). Direction of Improvement of Privacy Protection Regulation for Promoting Artificial Intelligence Industry. *Journal of Law and Economic Regulation, 9(2),* 161-176

[35] C. H. Lim, J. Y. Kim & J. H. Choi. (2009). Profiling of Cyber-crime by Psychological View, *Journal of the Korea Institute of Information Security and Cryptology, 19(4),* 115-124.

[36] EU. (2016). *General Data Protection Regulation (GDPR).*

[37] Personal Information Protection Commission & Korea Internet & Security Agency. (2020). *Personal Information Impact Assessment Performance Guide.*

[38] Cyber Security Strategic Headquarters. (2018). *Severity evaluation criteria (draft) for critical infrastructure service failures due to cyber attacks*(Online). https://www.nisc.go.jp/conference/cs/ciip/dai14/pdf/14shiryou 12-2 .pdf

[39] Korea Internet & Security Agency. (2021). Main contents of EU artificial intelligence (AI) regulation and protection of personal information. *Personal Information Protection Monthly Trend Analysis, 5*, 1-15

신 영 진(Young-Jin Shin) [정회원]

· 1996년 2월 : 성결대학교 행정학과 (행정학학사)
· 1998년 2월 : 단국대학교 일반대학원 행정학과(행정학석사)
· 2004년 2월 : 성균관대학교 일반대학원 행정학과(행정학박사)
· 2002년 9월 ~ 2014년 10월 : 성균관대 국제정보정책전자정부연구소 선임연구원
· 2004년 10월 ~ 2012년 7월 : 행정안전부 정보화전략실 전문위원
· 2012년 8월 ~ 2013년 2월 : 고려대학교 정보보호대학원 연구교수
· 2013년 3월 ~ 현재 : 배재대학교 AI소프트웨어공학부 정보보안학 부교수
· 관심분야 : 개인정보보호, 정보보호정책, 전자정부, 4차산업혁명 신기술
· E-Mail : jinsyj@yahoo.com