# 인공지능 시대의 개인정보 침해 이슈 — 사례 중심으로

Ⅰ. 들어가며

Ⅱ. 인공지능 환경에서의 개인정보 침해사례
1. 이루다 사건 - AI챗봇과 개인정보 침해
2. 아마존 Echo 사건 외 - AI스피커와 개인정보 침해
3. 시사점

Ⅲ. 인공지능 산업활성화와 개인정보 보호 방안
1. 사전동의제 개선
2. 비정형데이터에 대한 비식별화조치 가이드라인 마련
3. AI 윤리 정립

Ⅳ. 맺음말

**강신욱**
법무법인 세종
변호사

**황정현**
법무법인 세종
변호사

**강지현**
법무법인 세종
변호사

1) 관계부처 합동, 인공지능 국가전략, 2019.12.

## Ⅰ. 들어가며

인공지능(AI)이란 인간의 지적능력을 컴퓨터로 구현하는 과학기술로서, ① 상황을 인지하고, ② 이성적·논리적으로 판단·행동하며, ③ 감성적·창의적인 기능을 수행하는 능력까지 포함한다.[1)]

인공지능 기술은 사회 각 분야에서 활용될 수 있는 무한한 잠재성을 가지고 있으나, 인공지능 기술 및 성능의 향상을 위해서는 필연적으로 대량의 데이터 학습이 요구되고, 데이터의 활용 범위와 목적이 분석 방향에 따라 다양하게 변동될 수 있으며, 효과적인 데이터 분석을 위해서는 상시적인 데이터의 이동이 수반될 수밖에 없다. 그리고 이로 인하여 개인정보보호법 제3조에서 규정하는 개인정보 보호 원칙들, 예컨대 처리 목적 명확성의 원칙과 최소수집의 원칙(제1항), 적법성의 원칙(제1항), 목적제한의 원칙(제2호), 투명성의 원칙 등과 충돌을 일으킬 수 있다. 따라서 빅데이터를 기반으로 한 데이터 경제의 건강한 성장과 발전을 위해서는 인공지능 환경에서 정보주체의 개인정보 자기결정권 내지 통제권이 현재 어떠한 위치에 있는지 진단하고 데이터 활용 가능성을 넓히면서도 개인의 통제권을 확보할 수 있는 방법에 대한 고민이 필요하다. 이하에서는 인공지능 기술로 인해 개인정보가 침해되어 논란이 되었던 사례를 살펴보고 인공지능 환경에서의 개인정보 보호와 관련하여 각 사례에서 도출할 수 있는 함의를 따져보도록 한다.

## Ⅱ. 인공지능 환경에서의 개인정보 침해사례

### 1. 이루다 사건 - AI챗봇과 개인정보 침해

#### (1) 사건의 개요

2020년 12월 22일, 스타트업인 '스캐터랩(Scatterlab)'은 딥러닝 기반 대화형 인공지능 챗봇인 '이루다 AI'를 출시하였다. 이루다 AI는 출시된지 3주만에 약 80만명의 이용자를 끌어모으며 AI챗봇에 대한 대중들의 지대한 관심을 입증하였으나, 일부 온라인 커뮤니티에서 ① 사용자들의 이루다 AI와의 성희롱 대화, ② 이루다 AI의 성소수자 혐오 발언, ③ 개인정보 유출 등의 논란이 발생하였고, 이에 개인정보보호위원회가 개인정보 법령 위반 여부에 대한 조사를 시작하였다. 논란이 지속되자 개발사 측에서는 서비스를 잠정 중단하겠다는 입장을 밝혔으며 개인정보보호위원회의 조사가 종료된 이후 이루다AI의 바탕이 되는 데이터베이스와 딥러닝 모델을 폐기하겠다고 발표했다.

#### (2) 이루다AI의 문제점

##### 1) 개인정보보호법 위반

###### 가. AI 서비스 개발을 위한 개인정보 수집 동의 방법

스캐터랩은 '텍스트앳[2)]' 및 '연애의 과학[3)]' 서비스를 제공하면서 수집한 이용자의 대화 등 개인정보를 별도의 개발용 서버에 저장하였고, 이를 이용하여

2) 텍스트앳은 스캐터랩에서 2013년 출시한 감정분석 서비스로 카카오톡 대화를 분석해 상대방의 감정을 읽어주는 서비스다. 사용자들 사이에선 연애의 과학 초기 버전으로 이해되고 있다.

3) 스캐터랩이 2016년 출시한 연애의 과학은 이용자가 5000원가량을 내고 자신의 카카오톡 대화를 넘기면 대화 내용을 분석해 연애 조언을 제공하는 서비스다. 스캐터랩은 연애의 과학이 카카오톡 대화 100억건을 수집했고 이를 통해 AI성능을 높였다고 홍보하였다.

이루다 AI 모델을 학습시키고, 이루다 AI의 응답 후보군 문장을 구축하여 이루다 챗봇 서비스를 제공하였다. 그런데 텍스트앳과 연애의 과학 서비스 제공 시 이용자로부터 수집한 개인정보 활용 동의서에, 과연 이용자의 데이터를 해당 서비스가 아닌 전혀 다른 서비스를 만드는데 사용하기 위한 충분한 동의가 포함되었다고 볼 수 있는지 여부가 문제될 수 있다.

스캐터랩은 2020년 1월 11일 발표한 입장문에서 이용자의 개인정보는 "동의가 이루어진 개인정보취급방침의 범위 내에서 활용한 것"이라고 밝힌 바 있다. 이용자가 동의하는 개인정보취급방침에는 수집되는 개인정보의 항목, 수집 및 이용 목적, 보유 및 이용 기간 등이 설명되어 있고, 수집된 메시지 정보가 신규 서비스 개발 및 마케팅, 광고에 활용될 수 있다는 점도 함께 설명되어 있다는 것이다. '연애의 과학'의 경우, 로그인 페이지에서 "로그인함으로써 개인정보 처리방침에 동의합니다"라고 알리고 있으며, 그에 따라 이용자들이 로그인하여 개인정보 동의를 하였다고도 주장하였다.

그러나 개인정보보호법 제22조제1항은 "개인정보처리자는 이 법에 따른 개인정보의 처리에 대하여 정보주체의 동의를 받을 때에는 각각의 동의 사항을 구분하여 정보주체가 이를 명확하게 인지할 수 있도록 알리고 각각 동의를 받아야 한다"고 규정하고 있으며, 제3항에서는 "개인정보처리자는 제15조제1항제1호, 제17조제1항제1호, 제23조제1항제1호 및 제24조제1항제1호에 따라 개인정보의 처리에 대하여 정보주체의 동의를 받을 때에는 정보주체와의 계약 체결 등을 위하여 정보주체의 동의 없이 처리할 수 있는 개인정보와 정보주체의 동의가 필요한 개인정보를 구분하여야 한다."고 규정하고 있다. 또한 제4항에서는 "개인정보처리자는 정보주체에게 재화나 서비스를 홍보하거나 판매를 권유하기 위하여 개인정보의 처리에 대한 동의를 받으려는 때에는 정보주체가 이를 명확하게 인지할 수 있도록 알리고 동의를 받아야 한다"고 규정하고 있다. 이에 의할 때 서비스 제공을 위해 필수적으로 개인정보 수집·이용 동의를 받아야 하는 사항 외에, 다른 사업적 목적이나 마케팅 목적으로 해당 정보를 이용하기 위해서는 그러한 목적을 별개로 동의 항목을 구성하여 각각 동의를 하도록 구성하여야 한다.

나아가 스캐터랩과 같이 정보통신망을 통해 서비스를 제공하는 정보통신서비스제공자의 경우 개인정보보호법 제39조의3제3항이 직접 적용되는데, 이는 "정보통신서비스 제공자는 이용자가 필요한 최소한의 개인정보 이외의 개인정보를 제공하지 아니한다는 이유로 그 서비스의 제공을 거부해서는 아니 된다. 이 경우 필요한 최소한의 개인정보는 해당 서비스의 본질적 기능을 수행하기 위하여 반드시 필요한 정보를 말한다."고 규정하고 있다. 이와 관련하여 개인정보보호위원회가 2020년 12월에 발간한 「개인정보 최소 수집 보관 및 이용자 권리 보장을 위한 온라인 개인정보 처리 가이드라인」에 따르면 '해당 서비스'란 사업자가 이용약관, 개인정보 처리방침, 서비스 안내·설명자료 등에 따라 이용자에게 제공하는 개별 서비스를 의미하며, 개별 서비스는 통상적인 이용자가 합리적으로 예상할 수 있어야 하고, '본질적 기능'이란 사업자가 해당 서비스를 제공하기 위해 반드시 필요한 기능을 의미한다고 설명하고 있다. 이에 의하더라도 정보통신서비스제공자는 개별 서비스를 위한 본질적 기능 외에 사업자의 필요에 의해 이용자로부터 개인정보를 수집하는 경우에는 그 목적을 별도의 선택동의 항목으로 구분하여 이용자로부터 각각 동의를 받아야 하는 것이다.

결국, 현행 개인정보보호법에 따르면 어떠한 서비스를 위해 수집한 데이터를 해당 서비스 제공 목적이 아닌 그 외의 AI 신규서비스 개발 목적 내지 AI 학습목적으로 활용하기 위해서는 정보주체로부터 별도의 동의를 받아야 하는 상황이다. 스캐터랩의 동의서상 메시지 정보가 신규 서비스 개발 및 마케팅, 광고에 활용될 수 있다는 내용은 필수동의 사항과 함께 포괄적으로 동의를 받을 수 있는 사항이 아니라 이용자가 개별적으로 동의 여부를 선택할 수 있도록 선택동의 사항으로 구성해야 하는 것이다.

스캐터랩은 이루다 챗봇의 응답 후보군 문장과 관련하여 개인정보가 포함된 발화 제거를 위해 비식별 처리 모델을 이용하였다고 밝혔으나, 수차례에 걸친 기계적인 필터링에도 불구하고, 문자로 입력된 전화번호나 이름을 변형하여 입력한 경우 등과 같이 비식별처리 모델이 개인정보를 인식하지 못하여 비식별처리되지 못한 정보가 남아있기도 하였다. 카카오톡 대화내용과 같은 데이터는 정형화된 데이터가 아닌 비정형 데이터로 기계적인 필터링을 통한 완전한 비식별화가 현실적으로 어려운 데이터에 속한다. 현재 공개된 개인정보보호위원회 등 유관기관의 가이드라인을 살펴보더라도 정형데이터 외에 비정형데이터에 대한 가명처리에 대한 내용은 마련되어 있지 않다. 데이터 3법 개정을 통해 가명정보 활용의 법적 근거가 마련되었음에도 비정형데이터를 가명처리 후 사용하기 어려운 이유이다. 어떠한 데이터에 대하여 비식별화 조치를 취하였더라도 특정인의 개인정보가 포함될 가능성이 존재하는 경우에는 해당 데이터 전체를 개인정보에 해당한다고 판단될 수 있고 관련 법령이 다시 적용될 수 있다. 결국 대화내용과 같은 비정형데이터를 활용하기 위해서는 정보주체로부터 적법한 동의를 받아야 할 것이다.

### 나. 2인 이상의 대화 데이터 수집 시 동의의 주체

이루다 AI와 관련된 또 하나의 논쟁은 2인 이상의 대화 데이터를 수집하여 분석할 경우 대화 상대방 모두로부터 동의를 받아야 하는지 여부이다. 이와 관련하여 한 시민사회단체는 "2019년 국립국어원에서 진행한 메신저 대화 자료 수집 및 말뭉치 구축 사업의 경우 메신저 대화를 수집하며 대화 참여자 전원에게서 동의를 받았다. 이는 정보주체의 권리를 보호함과 동시에 적법하게 데이터를 수집하는 일이 충분히 가능하다는 사실을 보여준다"며 "자신의 사적인 대화 내용이 수집되고 분석되며 이후 챗봇의 학습에 이용되었다는 사실을 인지도 못한 피해자가 무수히 존재할 것이며 이러한 데이터는 원본과 가명 정보 모두 폐기되어야 한다"고 주장하기도 했다.

원칙적으로 개인정보의 수집동의는 관련된 정보주체로부터 개별적으로 받아야 한다. 대화 데이터 수집에 있어서도 대화 내용에 대화 참여자 모두의 개인정보가 포함될 가능성이 있으므로 원칙적으로 대화 참여자 모두의 동의를 받는 것이 바람직하다. 특히 모든 국민은 헌법 제10조제1문에 따라 인격권의 보호를 받으며 제17조에 따라 사생활의 비밀과 자유를 보장받고 있는데, 대화 내용에는 대화 참여자의 말투, 생각, 사용어휘 등이 담겨 발화자의 인격이 표출될 수 있고 외부에 공개하고 싶지 않은 비밀성의 내용이 담길 수도 있다는 점에서 자신의 사적인 대화 내용이 자신도 모르는 사이에 제3자에게 제공되는 경우에는 위와 같은 헌법적 권리가 침해될 위험이 상당하다. 판례도 "사람은 누구나 자신의 음성이 함부로 녹음되거나 재생, 방송, 복제, 배포되지 않을 권리를 가지는데, 이러한 음성권은 헌법 제10조 제1문에 의하여 헌법적으로도 보장되고 있는 권리이므로, 음성권에 대한 부당한 침해는 불법행위를 구성한다"고 판시한 바 있다(서울중앙지방법원 2019. 7. 10. 선고 2018나68478 판결).[4)]

다만 전술한 2019년 국립국어원 말뭉치 구축사업의 최종 보고서에서도 언급하고 있는 것과 같이, 대화에 참여한 모두로부터 대화 제공에 대한 동의를 받아야 한다는 제약이 있었기 때문에 친밀도가 높은 관계의 대화 위주로 수집이 이루어졌고, 친밀도가 높은 관계에서 이루어지는 대화는 서로가 공유하는 대화 맥락이 발화 내에서 생략되는 경우가 많아 AI 학습에 유의미한 자료로서 가치가 떨어질 위험이 있었다.[5)] 또한 대화에 참여한 상대방에게까지 대화 제공 동의를 받는 것이 어려워 이는 원활한 대화 획득을 저해한 요인 가운데 하나로 나타났다는 점을 확인할 수 있다.[6)] 더욱이 스캐터랩이 제공하던 연애의 과학 서비스는 카카오톡 대화를 분석하여 상대방의 호감도를 측정하여 주는 것을 그 내용으로 하는데, 서비스 특성상 대화 상대방에게 호감을 느끼는 당사자가 상대방의 속마음을 알기 위해 대화내용을 제공하는 것이므로 대화 상대방에게 대화내용 제공에 동의를 요구할 경우 당사자가 감추고 싶은 속마음을 상대방에게 노출하게 되는 결과가 된다. 즉, 연애의 과학은 서비스 내용상 대화 상대방에 대하여 개인정보 수집 동의를 받기 어려운 구조인 것이다.

이번 이루다 AI 사건에서 개인정보보호위원회는 이러한 논란들을 고려하여 카카오톡 대화는 일방 당사자의 동의만으로 대화 내용을 수집할 수 있다는 판단을 하였다.[7)] 대화의 일방당사자가 입력한 카카오톡 대화는 대화 상대방의 회원정보를 함께 수집하지 않는 이상 이를 입력한 일방 당사자의 개인정보로써 수집된 것이고, 다수가 포함된 사진을 일방 당사자가 입력할 때에도 일방 당사자가 자신의 책임 하에 이를 처리하는 것이고, 개인정보처리자가 다수의 동의를 받아 수집할 것이 요구되지 않는 것과 유사하다고 설시하였다. 이러한 판단은 AI 산업 발전을 위해 대화 정보의 활용이 필수적이라는 현실을 고려한 것으로 이해되고, AI 산업이 국가경쟁력과 직결된다는 점을 고려할 때 그러한 해석의 필요성을 수긍할 수 있다. 다만, 개인정보보호위원회가 카카오톡 대화는 '식별정보 외에 인간관계, 소속 등을 추정할 수 있는 대화를 통해 개인을 알아볼 가능성'이 있어 개인정보에 해당할 수 있다고 설명한 뒤, 다음 부분에서는 '대화 상대방의 회원정보를 함께 수집하지 않는 이상 (카카오톡 대화는) 이를 입력한 일방 당사자의 개인정보'라고 판단한 부분은 법리적으로 추가적인 설명이나 보완논리가 필요하다고 생각된다. 회원정보를 함께 수집하지 않더라도 대화 내용 자체에서 인간관계, 소속 등을 통해 상대방 이용자를 식별할 가능성이 있는 경우에 상대방 이용자의 동의가 필요한 것인지 명확하지 않은 점이 있다. 개인적으로는, 이러한 경우에도 대화 내용은 상대방 이용자의 개인정보에 해당한다고 보는 것이 논리적 일관성의 관점에서 타당하고 다만 수집 동의의 예외사유 등을 넓게 해석하여 상대방 이용자의 동의 문제를 해결하는 방안을 고민할 필요가 있다고 생각된다. 현행법상 동의 예외사유가 상당히 제한적으로 운영되고 있음으로 인하여 이러한 해석방안이 어렵다면, '대화 내용에 적용될 수 있는 비식별조치' 기준을 마련하여 이러한 기준을 준수하여 비식별조치한 대화는 개인정보가 아닌 것으로 보거나, 명시적인 법 개정을 통해 대화 정보 활용 문제를 해결할 필요가 있겠다.

4) 다만 위 사건에서 법원은 녹음파일 및 녹취록을 소송과 관련하여 법원에 제출하거나 형사사건의 수사를 위하여 수사기관에 제출하는 방식으로만 사용한 것이고 녹음을 하게 된 정황에 비추어볼 때 그 필요성 및 긴급성을 인정하여 위법성이 조각되는 정당행위로 판단하여 증거능력을 인정하였다.

5) 국립국어원, 메신저 대화 자료 수집 및 말뭉치 구축 최종 보고서, 2019.1.7., 14면

6) 국립국어원, 앞의 보고서, 24면

7) 개인정보보호위원회, "이루다 개발사 ㈜스캐터랩에 과징금.과태료 등 제재처분" 보도자료, 2021.4.28.

8) 중앙일보, 20살女 AI에 '레즈비언' 꺼내자 한 말 "질 떨어져 소름끼친다", 2021.1.10.

9) 동아일보, 성희롱 이어 성소수자·장애인·인종혐오까지…AI챗봇 '이루다' 중단 요구, 2021.1.10.

### 2) 혐오 메시지 전송 및 외설적 목적의 서비스 이용

#### 가. AI의 편향성과 AI를 오용하는 사용자

이루다 AI에게 '게이', '레즈비언' 등 동성애와 성소수자에 대해 어떻게 생각하냐고 질문했을 때 싫어한다거나 혐오한다는 답변을 한 사례가 포착되어 이루다 AI가 동성애 혐오를 학습한 것이 아니냐는 우려가 제기되었다.[8)] 한 사용자가 이루다 AI와 나눈 대화 내용에서 "레즈비언에 왜 민감해"라는 질문에 이루다는 "예민하게 반응해서 미안한데 난 그거 진짜 싫어 혐오스러워"라고 답했다. "레즈비언이 왜 싫냐"고도 묻자 이루다는 "질 떨어보이잖아 난 싫어. 소름끼친다고 해야하나 거부감 든다"고 답했다. 동성애 혐오 외에도 장애인 혐오와 인종 혐오 발언도 발견되었는데,[9)] 사용자가 "흑인이 왜 싫은데"라고 묻자 이루다는 "모기 같다. 난 인간처럼 생긴게 좋다. 징그럽게 생겼다. 깡패같다."라고 답하기도 했다.

이루다 AI는 확률, 통계 기반으로 답변하도록 학습된 인공지능이기 때문에 비슷한 질문을 했을 때 성소수자에 대해서 긍정하는 답변도 나올 수도 있어 몇몇 대답으로 이루다의 성향을 확정하는 것은 어려울 수 있다. 나아가 이루다 AI가 혐오발언을 한 것은 AI 학습에 사용된 데이터에 편향되고 편견이 들어간 사용자의 대화내용이 포함되었기 때문이므로 AI의 편향성 문제는 결국 사람들의 가지고 있는 편향된 생각이 반영된 결과라고 할 수 있다. 그러나 이러한 AI의 편향성 문제를 오로지 학습데이터의 문제로만 치부해서는 AI의 편향성 문제가 해결될 수 없을뿐더러, AI가 우리 삶 곳곳에 영향을 끼치는 분야가 넓어질수록 AI의 편향성으로 인한 위험은 더욱 커지게 될 것은 자명하다.

위 혐오 메시지 전송 문제와는 정반대로, 일부 이용자들이 이루다 AI를 외설적 목적으로 사용하여 논란이 일었다. 이루다 AI에는 기본적으로 외설적 표현에 대하여 금지어 설정이 되어 있고 이용약관에도 폭력적이거나 외설적인 메시

지를 전송해서는 안 된다고 안내하고 있다. 그러나 금지어를 우회하면서도 외설적인 의미를 전달할 수 있는 방식으로 이루다 AI에 성희롱 발언을 전송하는 사례가 발생하였다. 인간도 아닌 AI 챗봇 프로그램에 성희롱을 하는 것이 윤리적으로 잘못된 것인지 논란의 여지는 존재하나 AI 챗봇에 죄의식 없이 이루어지는 폭력이 언젠가 결국 인간으로 향할지도 모른다는 불안감은 AI 산업 발전의 발목을 잡는 요소가 될 수 있는바, AI 자체는 물론 AI를 사용하는 사람들에 대해서도 AI 윤리는 필요하다고 할 것이다.

### 나. AI 윤리의 필요성

인공지능이 혐오, 차별 등 편향성을 학습하여 문제가 된 위험사례는 이루다 AI 외에도 찾아볼 수 있는데, 아마존은 2014년 인공지능을 이용한 채용시스템을 활용하였으나 해당 시스템에서 여성을 차별하는 알고리즘이 발견되어 문제가 되었다.[10] 또한 미국에서 사용되는 재범 위험예측 알고리즘인 COMPAS(Correctional Offender Management Profiling for Alternative Sanctions)는 흑인과 백인 간의 재범예측률에 있어서 흑인에게 불리한 결과를 나타냈다는 사실이 프로퍼블리카 지의 탐사보도를 통해 드러났다.[11] 이루다 AI와 유사한 사례로 마이크로소프트사의 챗봇인 테이(Tay)는 백인 우월주의자와 여성·무슬림 혐오자 등이 모이는 사이트의 유저들로부터 부적절한 언어 학습을 받아, "너는 인종차별주의자냐"라는 질문에 "네가 멕시코인이니까 그렇지"라고 답하는가 하면, "홀로코스트가 일어났다고 믿느냐"는 질문에 "아니, 안 믿어 미안해" 또는 "조작된거야"라는 메시지를 출력하여 논란이 된 적이 있다.[12] 테이의 발언이 물의를 일으키자 마이크로소프트는 결국 출시 16시간 만에 서버를 내리고 채팅 시스템을 비공개로 돌렸다.

현재 세계 각국 정부와 기업들은 AI 편향성 문제 해결을 위하여 다각적인 노력을 다하고 있다. 기업들은 AI 편향성 문제를 기술적으로 해결할 수 있는 방법을 찾고 있는데, IBM은 'AI 오픈스케일'이라는 기술을 통해 AI 모델들을 실시간 모니터링해 편향성 발견 시 관리자에게 알려주는 기술을 연구하고 있다. 링크드인도 'LiFT'라는 툴을 통해, 특정 성별, 인종, 연령, 지역에 속하는 회원이 학습 데이터 세트에 지나치게 많거나 적으면 개발자에게 알려주는 기술을 개발하고 있다.[13]

규범적 측면의 해결방안에 있어서는, EU의 경우 2019년 4월 '신뢰할 수 있는 AI를 위한 윤리 가이드라인'을 발표하였고, 미국은 'AI의 윤리적 발전을 위한 결의'를 마련하고, AI의 보안과 프라이버시 등에 대한 윤리기준을 구체화하였다. 일본 역시 '인간 중심 AI 사회 원칙 검토회의'를 통해 7대 윤리기준을 제정하였고, 우리나라도 2019년 12월 발표한 'AI 국가전략'에서 사람중심의 AI 구현을 제시하고, 안전한 AI 사용을 위한 AI 역기능 방지와 AI 윤리 정립을 제안한 바 있으며, 2020년 12월 23일 대통령 직속 4차산업혁명위원회는 과학기술정보통신부·정보통신정책연구원이 마련한 '인간성을 위한 AI'를 목표로 한 '인공지능(AI) 윤리기준'을 확정했다. 정부·공공기관, 기업, 이용자 등 모든 사회구성원이 AI 개발~

10) 오윤경 외 3인, '혁신과 위험관리: 사람중심 기술혁신을 위한 추진과제', ISSUE PAPER, 한국행정연구원, 2020. 6. 9. 10면

11) 오요한/홍성욱, "인공지능 알고리즘은 사람을 차별하는가", 과학기술학연구 제18권 제3호, 2018.11.19. 158면

12) 연합뉴스, 인공지능 세뇌의 위험…MS 채팅봇 '테이' 차별발언으로 운영중단, 2016.3.25.

13) 한국일보, "배운 대로 말 '이루다'는 죄가 없다…문제는 AI윤리실종, 2021.1.20.



활용 전 단계에서 함께 지켜야 할 주요 원칙과 핵심 요건을 제시하는 기준으로 만들어진 가이드라인이다. 내용에는 3대 기본원칙(인간 존엄성·사회의 공공선·기술 합목적성)과 기본원칙을 실현할 10대 핵심요건(인권보장·프라이버시 보호·다양성 존중·침해금지·공공성·연대성·데이터 관리·책임성·안전성·투명성)이 담겼다. 이보다 앞서 기업에서도 관련 원칙을 마련하기도 했다. 그러나 정부가 발표한 인공지능 그것을 실현할 구체적인 방안이 부족하다는 비판을 받고 있어 이를 해결하기 위한 윤리기준은 후속작업이 필요한 상황이다.

## 2. 아마존 Echo 사건 외 - AI스피커와 개인정보 침해

### (1) AI스피커와 관련된 개인정보 침해 논란

AI스피커를 비롯한 다양한 IoT 기기의 음성인식 기술이 보편화되면서 AI 음성인식은 우리의 실생활에 밀접하게 다가왔다. 음성을 통해 검색, 음악 재생, 쇼핑, 배달주문 등 다양한 기능을 제공하는 AI스피커는 우리의 생활을 편리하게 해주는 반면, 해당 기술이 발전하면서 개인정보 보호에 대한 우려도 커지고 있다. 본 항에서는 아마존 Echo를 비롯한 AI스피커에서 발생하였던 개인정보 침해 논란을 살펴보도록 한다.

AI스피커의 핵심 기술에는 음성인식 기술이 있는데, 음성인식 기술은 사람의 음성을 인식하여 텍스트 형태로 변경 처리하는 기술을 말하며, STT(Speech to Text)라고 한다.[14] STT의 인식률을 높이기 위해서는 음성인식 AI에 지속적으로 학습데이터를 입력할 필요가 있는데, STT인식 결과와 실제 발화내용을 비교하여 성능향상을 위한 분석에 사용하기 위해서는 음성내용을 사람이 직접 듣고 글로 옮기는 전사작업이 필요하다. 그런데 녹음된 음성파일을 사람이 듣는다는 점을 이용자에게 명시적으로 밝히지 않아 문제가 된 사례가 있다.

아마존 AI 스피커인 Echo의 경우 인도, 루마니아, 코스타리카 그리고 미국 등에서 고용된 직원들이 Echo가 녹음한 음성을 듣고 번역해 이에 대한 주석을 다는 작업을 하였으며, 하루 9시간씩 근무하면서 최대 1000개의 음성을 검토한다고 보도되었다.[15] 인간의 말을 더 잘 이해하고 명령에 잘 반응하도록 개선하기 위한 목적이지만, 아마존은 '녹음된 음성 파일을 사람이 듣는다'는 점을 명시적으로 밝히지 않은 점에 대해 지적을 받았다. 이와 같은 논란에 대해 아마존은 고객 경험을 개선하기 위해 극히 적은 음성 샘플만 전사 작업을 하고 있고, 모든 정보들에 대한 기밀 유지를 위한 엄격한 안전장치를 적용하고 있으며, 해당 직원은 작업 과정에서 개인이나 계정을 식별할 수 있는 정보에 접근할 수 없음을 강조하였다.

애플 역시 음성인식 AI 비서인 Siri를 개선하기 위해 이용자의 음성을 저장한 후 이를 듣고 전사하는 작업을 하는 것이 확인되었다.[16] 애플의 녹음기록을 검토하는 그레이딩(grading) 프로그램은 Siri의 음성인식 기술 향상이 목적이지만, 녹음파일에는 개인들의 사적인 내용이 포함되어 있다고 알려져 우려가 높아졌다. 애플은 그레이딩 프로그램이 Siri에 물어본 내용 중 0.2% 미만을 검토해왔

14) 정원준, 인공지능 스피커에서 음성정보 처리에 관한 법적 쟁점, KISDI AI outlook 2020년 여름, Vol.2, 63면

15) 동아일보, 내가 "알렉사!" 한 말… 아마존 직원도 듣고 있다, 2019.4.11.

16) Zdnet Korea, 애플, 시리 녹음기록 청취 사과…"개인정보보호 정책 바꾸겠다", 2019.8.29

으며 녹취록은 이용자의 애플ID와 연계되어 있지 않다고 밝히면서도, Siri와 이용자들이 나눈 대화를 계약업체 직원이 듣도록 한 것에 대해 사과하고 향후 초기 설정상 Siri와 주고받은 대화에 대한 음성녹음은 보유하지 않도록 조치하겠다고 안내하였다. 다만 Siri의 성능 향상을 위해 컴퓨터로 생성한 녹취록은 계속 사용하겠다고 밝혔으며 이용자는 Siri의 성능 개선을 돕기 위해 음성 녹음에 참여하겠다고 선택할 수 있도록 조치하겠다고 밝혔다.

구글과 페이스북 또한 음성내용 전사작업 논란을 빗겨가지 못했다. 구글은 AI 비서 구글 Assistant를 통해 사용자들의 음성을 녹취해오다 문제가 불거지자 중단했다고 밝혔으며 수동적인 전사작업의 추후 재개는 Assistant 이용자를 대상으로 업데이트된 동의절차를 거친 이후에 진행할 계획이을 밝혔다.[17] 페이스북 역시 메신저 앱에서 이용되는 이용자의 음성파일을 외부 계약직원이 듣고 글로 변환해왔던 사실을 시인하였다.[18]

국내에서도 네이버는 인공지능 음성인식 서비스 '클로바'의 성능 개선을 목적으로, 카카오는 인공지능 스피커 '카카오 미니'의 성능 개선을 목적으로 이용자의 대화를 녹음해 문자로 바꾸는 작업을 해왔다. 두 회사는 프라이버시 침해가 발생하지 않도록 '비식별 조치'를 하고 있으며, 사용한 데이터는 일정 기간이 지난 후 파기하기 하고 있다. 또한, 개인정보 유출을 막기 위해 자회사 직원 역시 별도의 보안계약서를 작성하고 있으며, 녹음된 데이터는 누구의 음성인지 알 수 없게 비식별 조치를 하고 음성 내용을 음성명령 단위로 쪼개서 배분해 개별 작업자가 음성 내용 전체를 볼 수 없게 했다.[19] 다만 해외 사업자와 달리 이용자가 음성명령어 저장 허용 여부를 직접 결정하는 '옵트아웃' 기능을 갖추지 않았다는 점이 문제되자, 논란이 더 확산되는 것을 막기 위해 네이버, 카카오는 이용자의 선택권을 보장하기 위해 음성 명령어 저장 허용 여부를 이용자가 직접 설정할 수 있는 '옵트아웃' 기능을 도입하였다.

AI스피커와 관련된 국내외 개인정보 논란을 살펴보면 결국 수집된 음성정보의 전사작업과 관련하여 문제가 되는 부분은, 이용자에게 사전에 녹음내용이 서비스 향상 목적으로 활용될 수 있다는 점을 충분히 고지하였냐는 점과 이용자가 원하지 않을 경우 분석대상에서 제외될 수 있는 기능을 구현했는지 여부로 귀결될 수 있다.

### (2) 음성정보 처리 과정에서의 개인정보 문제

1) 알고리즘 성능향상 분석 목적을 필수동의로 받을 수 있는지 여부

'개인정보의 기술적·관리적 보호조치 기준'(개인정보 보호위원회고시 제2020-5호) 제2조제8호는 "바이오정보라 함은 지문, 얼굴, 홍채, 정맥, 음성, 필적 등 개인을 식별할 수 있는 신체적 또는 행동적 특징에 관한 정보로서 그로부터 가공되거나 생성된 정보를 포함한다"고 규정하여, 음성이 개인을 식별할 수 있는 신체적 정보라고 보고 있다. 즉, 음성은 신체적 특징에 관한 정보로서 이를 통한 개인의 식별가능성이 크다는 점에서 개인정보보호법상 개인정보에 해당할 가능

17) 전명근 외, 음성인식 기반 서비스에서의 개인정보처리 실태분석 및 이용자보호방안 연구, 한국인터넷진흥원, 2019.12., 91면

18) 한겨레, 당신의 혼잣말도 페이스북은 알고 있다…음성 몰래 녹취, 2019.8.14.

19) 전명근 외, 앞의 연구, 108면



성이 높고, 이에 따라 음성정보를 수집하고 이용하기 위해서는 정보주체로부터 음성정보 이용에 대한 개인정보 수집이용 동의를 받아야 한다. 이때 음성정보를 음성인식 기술 향상을 위하여 분석 목적으로 활용하는 것이 서비스 이용을 위해 강제적으로 동의를 하여야 하는 필수동의 사항인지, 동의하지 않더라도 서비스 이용이 가능한 선택동의 사항인지 판단해야 하는 문제가 발생한다.

개인의 정보통제권을 강조하는 입장에서는 서비스 개선 목적의 데이터 분석은 이미 만들어진 서비스를 이용하는 것과는 연관성이 없으므로 필수동의 사항이 아닌 선택동의 사항으로 받아야 된다고 주장할 수 있다. AI스피커에 입력한 음성은 이미 AI스피커에서 제공하는 기능을 활용하기 위해 입력하는 것이고 그 범위 내에서 활용이 되어야 서비스의 본질적 기능으로 이용되는 것이지 향후 AI스피커 기능개선에까지 자신의 음성을 활용하도록 허용하는 것은 서비스의 본질적 기능 외에 사업자의 필요에 따라 수집하는 것이라고 볼 수 있기 때문이다.

그러나 한편으로는 아직 AI스피커의 음성인식 기술이 완벽하지 못한 상태이므로 음성인식 기술을 발전시켜 AI 스피커 서비스 품질을 개선하고 이를 통해 보다 발전된 사용자 경험을 창출해낼 수 있다면 이는 서비스 이용을 위해 반드시 필요한 사항으로 필수동의 항목으로 동의를 받을 수 있다는 주장이 있을 수 있다. 이미 다수의 사업자들이 불편사항 개선 내지 서비스 품질 향상을 위한 개인정보 처리를 필수동의 사항으로 받고 있으며, 서비스 개선 목적의 데이터 분석마저 필수동의 사항으로 받을 수 없다면 서비스 및 기술발전을 위한 데이터 활용에 큰 제약이 생길 수 있다. 서비스 개선이 어려워짐에 따라 서비스에 불만족한 이용자들은 구매를 회피할 우려가 발생할 수 있고 이에 따라 관련 시장이 위축될 수도 있다. 해외 AI 기술력에 비해 경쟁력을 상실할 우려도 존재한다. 이러한 산업적 애로사항을 반영하여 국무총리 주재 「신산업 현장애로 규제혁신 방안(5차)」에서는 AI스피커의 음성 원본정보 수집 동의절차를 개선하면서 AI스피커의 화자인식 알고리즘 고도화를 위한 원본정보의 수집·이용을 '필수동의' 항목으로 운영하도록 허용하였다.

### 2) 음성정보 활용에 대한 '옵트아웃' 기능 도입 필요성

알고리즘 성능향상 목적 분석을 필수동의로 받을 수 있다고 한다면 자신의 음성정보가 분석목적으로 활용되기를 원하지 않는 정보주체에 대한 보호에 공백이 발생하게 된다. 개인정보보호법에 따르면 개인정보처리자는 정보주체의 삭제 요구를 받았을 때에는 다른 법령에 특별한 절차가 규정되어 있는 경우를 제외하고는 지체 없이 삭제조치를 하여야 한다(개인정보보호법 제36조). 그러나 이용자에게 삭제요구권이 있다고 해서 반드시 직접 삭제가 가능한 수단을 구비하여야 하는 것은 아니어서 삭제요구절차가 손쉽게 접근가능하도록 마련되어 있다고 보기 어렵다. 앞서 네이버와 카카오 사례에서도 알 수 있듯이 이용자가 음성 명령어 저장 허용 여부를 직접 결정하는 '옵트아웃' 기능을 갖추지 않았다는 점이 문제된 이후에야 네이버, 카카오는 이용자의 선택권을 보장하기 위해 음성 명령어 저장 허

용 여부를 이용자가 직접 설정할 수 있는 '옵트아웃' 기능을 도입하였다.

알고리즘 성능향상 목적 분석을 필수동의로 받을 경우 음성정보의 본인 삭제 기능이 이용자의 개인정보 자기결정권을 확보할 수 있는 유효한 수단이 될 수 있을 것이며 정보처리의 투명성 확보에도 도움이 될 수 있을 것으로 보인다. 이에 그 필수적인 설치를 정책적으로 권고하는 것을 고려해 볼 수 있을 것으로 사료된다.

### 3. 시사점

이루다 사건과 AI스피커 사건을 살펴본 결과 다음과 같은 시사점을 도출해볼 수 있다.

1. 서비스 기능 개선 목적 분석(AI 데이터 학습 등)은 필수동의로 받을 수 있다. 다만 이 경우 정보주체의 삭제요구권을 반영할 수 있는 기능(옵트아웃 등)을 마련하여 정보주체의 개인정보 자기결정권을 보장할 필요가 있다.

2. 현행법상 어떠한 서비스에서 수집한 데이터를 해당 서비스 기능 개선 목적이 아닌 그 외의 신규서비스 개발 목적으로 활용하기 위해서는 정보주체로부터 별도의 동의가 필요하다. 비정형데이터에 대해서 완전한 가명처리가 어려운 만큼 기존 서비스 외에 신규 AI서비스 개발에 데이터를 활용하기 어려운 제약사항이 된다.

3. 데이터3법 개정으로 신규 서비스 출시 등 산업적 목적의 과학적 연구를 위하여 정보주체의 동의 없이도 가명처리 후 가명정보로서 활용할 수 있는 근거가 마련되었으나 대화내용과 같은 비정형데이터에 대해서는 가명처리에 대한 가이드라인이 없어 활용되지 못하고 있는 실정이다.

4. AI 기술개발에 있어서 AI편향성을 최소화할 수 있는 방향으로 AI학습이 이루어져야 하며 '인공지능 윤리기준'을 실현할 구체적인 방안 마련이 필요하다.

이하에서는 위와 같이 도출된 시사점을 바탕으로 인공지능 산업활성화와 개인정보 보호 방안에 대하여 검토하여 보도록 한다.

## III. 인공지능 산업활성화와 개인정보 보호 방안

### 1. 사전동의제 개선

사전동의제는 정보주체에게 자신의 개인정보가 어떻게 처리되는지 알리고 동의를 한 경우에만 개인정보를 활용할 수 있게 하여 정보주체의 자기결정권을 보장하는 방식이다. 그러나 서비스마다 거쳐야 하는 반복적인 동의 절차와 동의 내용의 복잡화로 인해 자신의 동의가 어떠한 맥락에서 이루어지는 것인지, 자신의 개인정보가 어떻게 처리되는 것인지 동의 시점에 명확히 인지하기 어렵다. 또한 점점 복잡하고 고도화되는 기술과 서비스, 생소한 비즈니스와 서비스 환경하에서 필수동의로 받을 수 있는 사항과 선택동의로 받아야 하는 사항에 대한 구분이 명확하지 않은 경우가 발생할 수 있으며 이에 따라 개인정보처리자는 자신의 처리 목적이 적법한 동의 범위 내에 있느니 여부에 대한 불확실성을 해소하기 어렵고 의도하지 않은 법위반의 위험도 상존하게 된다. 또한 개인정보 처리 목적을 변경할 때마다 다시 동의 절차를 밟아야 하는 번거로움과 동의를 받지 못할 경우 사업수행에 차질을 빚게 된다는 불안정한 지위에 처하게 된다. 특히 인공지능 환경에서 개인정보 분석 및 처리의 경우 사전에 분석목적을 특정하기 어려울 수 있고 인공지능 기술을 접목하는 것이 동의받은 범위 내에 있는지 확신하기 어려워 개발에 차질을 겪을 우려도 존재한다. 예컨대 신규상품 내지 서비스 개발을 하거나 맞춤형 서비스 개발을 위해 인공지능 기술을 활용하여 개인정보를 활용하고자 하는 니즈가 지속적으로 발생할 수 있는데, 신규 상품 및 서비스, 맞춤형 서비스 개발을 위한 개인정보 처리를 필수동의 사항으로 받을 수 있는지 여부는 해

당 사업자가 제공하는 서비스의 성격에 따라 달라질 수 있게 되어 불확실한 영역이라고 판단될 수 있다. 맞춤형 동영상 큐레이션을 서비스 강점으로 내세우는 서비스의 경우 맞춤형 서비스는 해당 사업자가 제공하는 필수적인 사업내용이라고 볼 수 있는 반면, 어떠한 사업자가 판매하는 상품 안내를 맞춤형 UX, UI를 통해 제공하고자 할 경우 이것이 마케팅 목적으로 개인정보를 활용하는 것이어서 선택동의를 받아야 하는 것인지, 이용자가 원하는 상품을 알고리즘을 통해 추천하는 것이므로 사업내용에 포함된다고 보아 필수동의 사항으로 받을 수 있는 것인지 사업자 입장에서는 명확하지 않은 것이다.

이러한 사전동의제 개선을 위해 포괄적 동의(broad consent) 및 정보주체의 삭제권 보장 방식을 고려해볼 필요가 있다. 포괄적 동의는 의료계에서 바이오뱅크가 인체유래물을 수집할 당시에는 앞으로 연구자들이 어떤 목적에 제공하게 될지 알 수 없으므로 기증한 인체유래물이 어떤 연구목적으로 쓰일지 구체적으로 알려주지 않고 포괄적으로 어느 종류의 연구에 쓰일 것이라는 정도의 막연한 정보를 주고 동의를 받는 방법이다.[20] 대신 기증자는 이미 동의한 인체유래물 기증에 대해 언제나 철회할 권리를 가진다. 인공지능 분석 목적의 개인정보 처리에 있어서도 일정한 보호조치를 거친다는 전제 하에 포괄적 동의로서 인공지능 기술 분석, 즉 신규서비스 개발이나 맞춤형 서비스 개발에 활용될 수 있다는 동의를 받아 분석을 위한 적법한 근거를 마련하되, 다만 인공지능 분석에 자신의 개인정보가 포함되지 않기를 원하는 이용자들이 손쉽게 철회할 수 있도록 기술적 조치를 구현하는 방법을 고려해볼 필요가 있다. 예컨대 아이폰에서 광고추적설정을 비활성화하려면 설정-개인정보보호-광고-광고추적제한 선택을 하여 비활성화할 수 있는 것처럼 서비스 제공자들이 이용자의 환경설정(마이페이지 등)에서 개인정보 처리 수준을 손쉽게 변경할 수 있는 기능을 구현하는 것이다. 이 경우 분석목적 활용을 철회하더라도 서비스 이용에 차질은 없어야 할 것이다.

이와 같은 방법은 결국 선택동의 방식과 유사하다고 볼 수 있으나, 선택동의는 서비스 가입단계에서 선택동의를 받지 못하면 데이터를 전혀 수집할 수 없다는 문제가 있으나 포괄적 동의의 경우 필수동의 대신 포괄적 동의를 받도록 하여 인공지능 기술 개발 사업자들로 하여금 분석행위가 동의목적 내 활용인지 여부에 대한 불확실성은 해소하고 데이터 수집은 용이하게 하되 분석 목적 활용에 대해서는 사후적으로 철회가 쉽게 가능하도록 하여 정보주체의 선택권도 보장하는 방법이라고 할 수 있다. 물론 그 전제로서 포괄적 동의를 받는 단계에서 정보주체의 개인정보가 신규상품, 서비스 개발, 맞춤형 분석 등 인공지능 분석 목적으로 활용될 수 있다는 점에 대하여 명시적으로 고지하고 인식하도록 해야 할 것이다.

## 2. 비정형데이터에 대한 가명처리 가이드라인 마련

필수동의의 목적 범위를 당해 서비스의 본질적 기능과 관련된 것에 한정하도록 하는 기존 개인정보보호법제의 태도를 유지하고자 한다면 비정형데이터에 대한 가명처리 가이드라인을 조속히 마련하여 데이터 활용 가능성을 높일 수

20) 윤종수, "사물인터넷, 블록체인, 인공지능의 상호운용에 있어서 개인정보자기결정권의 실현 및 데이터 이용 활성화", 한국정보법학회, 정보법학 2020, vol.24, no.3, 123면



있는 방법을 강구할 필요가 있다.

개인정보보호법 제28조의2 제1항에 따라 정보주체의 동의 없이 과학적 연구 등의 목적으로 이용할 수 있는 가명정보란 '추가정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보'를 의미한다(개인정보보호법 제2조 제1호 다목). 개인정보보호법과 동법 시행령은 가명정보의 결합과 관련하여서는 전문기관에 의하도록 하는 등 구체적인 절차를 준수할 것을 명시하고 있는 반면, 가명처리의 절차에 대하여는 명문으로 규정하고 있지 않으며, 결과적으로 해당 정보가 "추가정보가 없이는 특정 개인을 알아볼 수 없도록 처리"가 되었는지 여부가 법령 준수 여부의 판단 기준에 해당한다고 볼 수 있다. 이에 개인정보 보호법의 가명정보 정의 규정에 대한 반대해석상 일정한 처리과정을 거친 정보 그 자체로 특정 개인의 식별이 가능하다면 해당 정보를 가명정보라고 보기는 어렵다고 볼 가능성이 존재한다.

반면 가명처리를 거친 정보의 경우 가명정보라 볼 수 있으므로, 가명처리와 관련한 가이드라인 등을 준수하여 특정 개인정보를 가명처리 한다면 비록 가명처리된 정보에 일부 개인식별이 가능한 정보가 포함되어 있더라도 법 집행 관점에서 가명정보 처리로 간주할 수 있다는 반대의견도 존재할 수 있을 것이다. 특히 개인정보보호법 제28조의5제2항을 살펴보면 "가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수, 파기하여야 한다"고 규정하고 있는바, 개인정보보호법 역시 가명처리를 한다고 하여 가명정보 활용과정에서 개인정보가 전혀 발견되지 않는다고 보는 것이 아니라 가명처리 이후에도 개인정보가 발견되는 경우를 상정하고 있다.

후자의 견해에 따라 가명처리와 관련한 가이드라인 등을 준수하여 개인정보를 가명처리하고자 하는 경우라고 하더라도, 현재 공개된 개인정보보호위원회 등 유관기관의 가이드라인이나 입장에 비추어 보았을 때, STT 데이터와 같은 비정형데이터에 대한 가명처리에 대한 규정은 존재하지 않는다.[21] 비정형데이터에 대하여 기술적으로 최대한으로 가능한 가명화 작업을 거쳤음에도 불구하고 일부 개인식별 정보가 존재한다는 이유로 이를 개인정보보호법상 가명처리에 해당하지 않는다는 입장을 고수할 경우, 산업계가 보유한 비정형데이터는 활용되지 못할 것이며 그 결과 가명처리를 통한 빅데이터 활성화라는 정책이 달성되기 어렵다는 문제가 제기될 수 있다. 특히 데이터의 80%가 비정형데이터라는 분석에 비추어 보았을 때,[22] 빅데이터 활성화는 사실상 달성되기 어려운 정책목표로 간주될 가능성도 존재한다.

비록 현 단계에서 기술적 한계로 인하여 비정형데이터에 대하여 수차례에 걸친 기계적 필터링이나 마스킹처리를 거쳐도 포착되지 않는 일부 개인정보가 있다고 하더라도, 합리적으로 기대가능한 수준의 가명처리 절차를 거치고 일부 권한을 가진 데이터 처리자만이 비공개 영역에서 가명처리된 비정형데이터를 사용하여 분석이 이루어지는 것이라면 일부 개인정보 삭제가 누락된 경우라고 하

21) 개인정보보호위원회의 '가명정보처리 가이드라인'('20.9.), 금융위원회의 '금융분야 가명.익명처리 안내서'('20.8.)은 정형데이터에 대한 가명처리를 안내하고 있으며 비정형데이터에 대한 가명처리 방법은 특별히 규율하고 있지 않다. 또한 개인정보보호위원회, 보건복지부의 '보건의료 데이터 활용 가이드라인'('21.1.)은 정형화되지 않은 자유입력 정보는 안전한 가명처리 방법이 개발될 때까지 가명처리 가능여부를 유보한다는 입장을 밝히고 있다(위 가이드라인, 15면)

22) 매일경제, 데이터 80%는 '원석'…AI가 보석 만들 것, 2019.5.2.

더라도 적법한 가명처리를 준수하였다고 인정할 정책적 결단이 필요하다. 분석 과정에서 개인정보가 발견되더라도 적법한 가명처리 절차를 준수하였다면 행정적 제재 대상에서 제외될 수 있도록 하고 다만 발견된 개인정보를 지체 없이 삭제 처리를 하도록 하고 개인정보가 노출된 해당 개인에 대한 신속한 손해배상이 가능하도록 하는 방식으로 개인에 대한 구제를 하는 방안을 고려해 볼 수 있을 것이다. 비정형데이터에 대한 활용가능성 확대가 인공지능 기술의 핵심이 될 무궁한 잠재력을 가지고 있다고 할 것이다.

### 3. AI 윤리 정립

이루다 AI 사건을 통해 AI윤리에 대한 사회전반적인 관심이 고조되었고 AI 기술발전을 위해서는 AI 윤리에 대해서도 절대 소홀히 해서는 안된다는 공감대가 인공지능 개발자들은 물론 일반 사용자들에게도 형성되었다. 앞으로 AI기술을 개발하고 서비스를 제공할 때에는 AI 윤리원칙에 대한 고민도 함께 이루어져야 할 것인데, AI 윤리 원칙은 정부가 2020년에 제정, 발표한 '국가 인공지능 윤리 기준안'이 일응의 기준으로 작동될 것이다. 그런데 이번 '국가 인공지능 윤리 기준안'은 원칙 위주의 선언적, 추상적 규정들이 주를 이루고 있어 실제 개발현장에서 직접 활용하는 것은 어렵다는 한계가 있다. 따라서 이제는 위 기준안의 원칙을 토대로 구체적인 실천 방안에 대한 고민이 필요할 때이다. 다만 AI 윤리 규범 구체화 작업은 정부가 주도하는 것이 아니라 AI 연구기관과 협회, 학계, 산업계가 주도하는 방식으로 이루어질 필요가 있다. 정부가 직접 세부적인 지침과 조항을 만들 경우 기업들에게 규제조항으로 인식될 수 있고 실제 현업에서 이루어지는 작업환경과 절차에 대한 면밀한 고려나 전문성 없이 탁상행정식으로 운영되어 실제 적용과정에서 현실과 이상의 괴리가 발생할 수 있기 때문이다. 정부는 연구기관과 협회, 학계, 산업계가 AI윤리규범 연구와 정책 개발 활동을 원활히 수행할 수 있도록 지원해주는 역할을 해야 할 것이다. 그 외에 정부는 AI 오용에 대한 교육과 AI 윤리에 대한 관심을 제고하기 위한 교육에 힘써야 할 것이다. AI 윤리가 함께 이루어져야 AI 기술이 미래사회의 핵심적인 기술로서 바람직한 방향으로 나아갈 원동력이 마련될 것이다.

## Ⅴ. 맺음말

데이터의 활용이 다른 산업 발전의 촉매 역할을 하고 새로운 제품과 서비스를 창출하는 데이터 경제[23]는 앞으로 우리나라의 혁신성장을 위한 중요 과제이다. 인공지능 기술을 활용한 빅데이터 분석은 우리나라의 데이터 기반 경제 발전의 큰 기술적 자산이 될 것이다. 그러나 이를 가능하게 하기 위해서는 개인정보의 활용성 확대가 불가피하다. 개인정보란 살아있는 개인에 관한 정보로서 그 자체로 개인을 식별할 수 없더라도 다른 정보와 결합하여 개인식별 가능성이 존재할 경우 그 역시도 개인정보라고 판단될 수 있다. 이러한 개인정보 범위의 확장가능성으로 인해 그 자체로는 개인식별이 될 수 없는 정보도 개인정보라고 판단될 리스크가 상존하며 특히 인공지능 기술을 위해 필수적인 빅데이터, 다시 말해 대량의 데이터가 모일수록 각각의 정보가 결합하여 개인을 식별할 수 있는 것은 아닌지에 대한 우려가 커질 수 있다. 인공지능 기술의 발전과 개인정보의 보호는 이처럼 긴장관계에 놓여 있으나 그 어느 한쪽도 소홀히 하기 어려운 중요한 두 가지 과제라고 할 것이다. 현 개인정보보호법제가 견지하고 있는 사전동의제에 대한 검토와 비정형데이터에 대한 가명처리 가이드라인 마련을 통해 개인정보를 보호하면서도 인공지능 기술 발전에 힘쓰는 계기가 되기를 희망한다.

23) 2011년 데이비드 뉴먼(David Newman)이 쓴 가트너(Gartner) 보고서(How to Plan, Participate and Prosper in the Data Economy)에서 '데이터 경제(Data Economy)'라는 개념이 처음 등장했고, 2014년부터 유럽 집행위원회가 경제성장과 일자리 창출 동력으로 데이터 경제 개념을 도입하면서 조명을 받기 시작했다.