한국의료윤리학회지 제22권 제3호(통권 제60호) : 255-273 ⓒ한국의료윤리학회, 2019년 9월
Korean J Med Ethics 22(3) : 255-273 ⓒ The Korean Society for Medical Ethics, September 2019
pISSN 2005-8284 eISSN 2234-3598

투고일: 2019년 5월 6일, 심사일: 2019년 5월 22일, 게재확정일: 2019년 7월 4일

인공지능과 데이터 윤리에 관한 소고(小考): 인공지능에 활용되는 헬스 데이터를 중심으로

박미정*

요약

다양한 방식과 형태로 수집되는 개인의 헬스 데이터는 과학적 연구목적으로 이차 활용된다. 국내외 개인정보보호 법률은 과학적 연구를 위한 데이터 처리의 법적 근거로 정보 주체의 동의를 얻는 것 이외에 데이터의 익명성을 요구한다. 공공이익과 관련된 연구라면 정보 주체의 동의가 면제될 수 있는데, 이러한 양립 가능성을 적용하기 위해서는 법 제도 이외의 측면도 고려해야 한다. 본 논문의 연구 질문은 인공지능을 학습시키는 데 데이터가 활용된다는 점에서 시작한다. 인공지능 알고리즘의 특성을 우선 살펴보고 이러한 기술에 활용되는 헬스 데이터의 사례와 제기되는 문제점을 고찰하였다. 인공지능 기술 특성의 틀 안에 갇혀있는 문제는 윤리 논의의 주제가 될 수 있다. 법률을 통해 준비하고 해석하기 어려운 문제점으로 사전동의와 책임, 데이터 익명성에 대한 신화, 위험점수와 알고리즘적 차별을 설명하였다. 그리고 문제해결을 위해 변화를 꾀한 국내외 법률의 데이터처리 원칙에 대한 입법 과정과 법률 조항을 분석하여 해결을 위한 단서를 찾아보았다. 결론에서 '독(毒)'을 제거하는 데이터 처리, 사전적 절차로서의 법 제도와 조화, 사후적 판단을 위한 알고리즘의 투명한 설계를 제언하였다.

색인어

인공지능, 알고리즘, 익명과 가명, 윤리, 헬스

교신저자: 박미정. 서울대학교 의과대학 건강사회교육센터. Tel: 02-3668-7354. Fax: 02-745-8195. e-mail: apotre@snu.ac.kr
ORCID: https://orcid.org/0000-0002-3295-8649

* 서울대학교 의과대학 건강사회교육센터 전문위원

I. 시작하며

OECD의 인공지능전문가 그룹(OECD's AI Experts Group, AIGO)에 의하면 인공지능 시스템이란 인간이 부여한 목표에 따라 실제 또는 가상환경에 영향을 미치는 예측(predict), 권고(recommend), 또는 결정(decide)을 내릴 수 있는 기계 기반 시스템이라고 정의할 수 있다[1]. 대부분의 인공지능 응용프로그램은 지능형 결정을 배우고 지능적으로 결정하기 위해 엄청난 양의 이용 가능한 모든 데이터가 필요하다. 개인의 헬스 데이터는[1] 질병 치료 이외에도 연구와 교육에 이차 활용되는 경우가 많다. 헬스 데이터는 대부분 의료기관 내에서 다양한 종류의 기기와 시스템을 통해 생성되거나 수집될 뿐 아니라 의료기기에 따라 병원 밖에서도 병원으로 연결되는 데이터의 종류도 늘어나고 있다. 정보통신기술의 발전으로 연결성이 향상된 데이터는 인공지능을 학습시키는 기초데이터로써 활용될 수 있다.

인공지능 알고리즘은 많은 양의 데이터와 분리해서 생각할 수 없고, 데이터가 가진 잠재력이 실현되는 것은 알고리즘을 통해서이다[2]. 데이터 간에는 수많은 상관관계가 존재한다. 내 소비 패턴, 학습 패턴, 인터넷 검색 패턴, 병원 방문 패턴, 전화 이용 패턴 등이 종합되면 나 자신도 모르는 나에 대한 많은 상관관계가 생길 수 있다. 이러한 상관관계는 사람의 능력을 뛰어넘는 속도와 규모로 반복적이고 복잡한 패턴 매칭을 통해 도출된다. 여기에 패턴으로부터 스스로 배우는 딥러닝[2] 알고리즘이 결합하면 내 행동에 대해 예측이 가능해진다[3]. 이러한 패턴과 예측은 개인이나 어떤 특정한 집단에게도 적용될 수 있고, 장기간 추적 관찰하여 더 많은 상관관계를 집중적으로 밝혀낼 수도 있다.

데이터 분석에 기초한 프로파일링(profiling)은 개인정보 보호 이슈를 넘어 데이터 프라이버시 논의를 촉발했다[4]. '혼자 있을 권리(right to be let alone)'로 요약되는 전통적인 프라이버시 개념과 데이터 프라이버시라는 개념은 다르며 법률적인 차원에서도 새롭게 자리 잡았다. 유럽연합(European Union)이 2018년 5월 시행한 일반개인정보보호규정(General Data Protection Regulation, GDPR) [5]의 서문(recital)에는 정보 주체가 갖는 새로운 권리를 명시하고 있다. 정보 주체의 업무능력, 경제적 상황, 건강, 개인의 성향이나 관심사, 신뢰 또는 행동, 위치 또는 움직임과 관련된 측면을 분석하고 예측하며, 정보 주체에게 법적인 영향이나 이에 상응하는 중대한 영향을 미치는 결정에 적용받지 않을 권리가 명시되어 있다(GDPR recital 71). 이러한 권리에 대한 영향을 살피는 데이터 프라이버시는 자동화된 의사결정 도구로 인해 그 권리가 침해될 수 있고 복잡한 알고리즘에 활용되는 데이터는 사람들에게 실제적인 영향을 미칠 수 있다는 인식과 윤리적 논의의 대상이라는 것을 암시한다[6].

개인정보를 어떻게 처리할 것인가와 관련된 문제는 어떠한 법적 근거를 적용하여 누가 책임을 져야 하는지와 일정한 패턴으로 어떤 권리가 훼손되는지 그 영향을 평가하여 대안을 마련할 수 있다. 하지만 그러한 평가를 하는 일은 절대 쉽지

1) 본 논문에 사용된 헬스 데이터 개념의 범위는 데이터 생성과정과 처리목적에 따라 보건의료 정보를 지칭할 수 있고, 데이터의 형식에 따라 진료기록을 포함할 수 있고, 감지 장치(sensor) 등을 통해 수집되는 생리학적 데이터와 생체정보, 유전정보도 포함한다.

2) 본 논문에서 다루는 인공지능기술은 딥러닝으로 상정한다. 딥러닝은 기계학습(machine learning)의 최신 형태라고 할 수 있다. 딥러닝은 학습기반 접근방식이다. 특정한 문제 집합 내에서 확률적으로 옳을 것이라고 예상되는, 가장 넓은 범위를 포괄할 수 있는 규칙을 훈련을 통해 터득하게 된다. 기계학습 이전의 인공지능기술은 각각의 상황에 대한 규칙을 사전에 부여하고 그에 따라 판단하도록 하는 광의의 '전문가 시스템(expert system)' 중심이다. 전문가 시스템은 기반 지식과 추론 과정이 분리되어 있음으로 기반 지식의 변경에 따른 업데이트가 쉬우며, 판단과정의 해설이 쉽다. 반면 딥러닝은 설명이 불가해한 불투명성을 가지고 있다.

않으며 부적절하게 평가될 수도 있다. 예컨대 공공이익과 관련된 연구라면 정보 주체의 동의가 면제될 수 있는데 연구 가치에 대한 인정과 평가는 누가 어떻게 해야 하는지와 같은 양립 가능성의 논리를 펼치기 위해서는 기술적 대안으로는 부족하다. 익명성이 확보된 데이터를 연구에 활용하도록 허용한다면 데이터의 익명화 방법으로 달성하기 어려운 프라이버시 보호 문제는 없는지 평가하는 일은 법적 범위를 넘어서는 것이다[7].

본 논문에서는 인공지능 알고리즘의 특성을 우선 살펴보고 이러한 기술에 활용되는 헬스 데이터의 사례와 제기되는 우려를 고찰하였다. 그리고 문제점을 해결하기 위해 변화가 필요했던 국내외의 법률 내용과 입법 과정을 분석하였다. 나아가 인공지능 알고리즘과 헬스 데이터의 특성으로 말미암는 윤리적 문제와 고려사항을 찾아보고 데이터 윤리에서 그 해답을 모색해보았다.

II. 의료에서 인공지능에 활용되는 데이터의 요건과 위험

의료 분야에서 인공지능의 활용은 빠르게 발전하고 있다. 현재 사용되고 있거나 개발 중인 많은 응용 프로그램은 대부분 특정 의료기술이나 약품 사용의 결과를 예측하고 보다 효율적인 환자 관리를 위해 이용된다. 다양한 요인들의 복잡한 상호작용에 대한 확률을 분석해서 의료오류를 감소시키고 환자 스스로 증상관리나 만성질환 관리를 할 수 있도록 돕는 등 지속해서 환자 중심의 건강관리를 할 수 있다는 가능성을 강조한다. 국내외 개인정보보호와 관련된 법률은 개인정보수집과 처리에 관한 원칙을 가지고 있다. 최소수집 원칙이 있고, 헬스 데이터를 민감 정보의 범주에 속하는 데이터로 구분한다면, 수집부터 활용까지 충분한 설명에 의한 동의가 필요하다. 그 활용의 정당성도 엄격한 심의윤리 기준을 적용하여 확보되어야 한다. 공공기관에서 수집 · 관리되는 건강보험과 관련된 데이터는 개방형 시스템을 통해 특정 데이터베이스를 사용할 수 있고, 연구목적으로 활용되는 특정 데이터는 계약을 통해 받을 수도 있다. 연구목적으로 활용되는 데이터를 보호하기 위한 법률의 요건은 상관관계를 찾기 위해 필요한 데이터를 미리 선택하고, 이용목적을 정의하고, 정보 주체의 동의를 구하거나 개인정보처리를 제한하는 방식을 취한다.[3)]

문제는 직접 환자치료에 사용되지 않는 정보라도 데이터베이스로 수집 · 보관되어 있어서 연결된 다른 시스템을 통해 공유되거나 데이터가 이동된 후라면 정보 주체의 권리행사는 어려워 질 수 있다. 또한 데이터에 오류가 있으면 이를 찾아서 완벽히 제거하기가 어려우며 데이터 오류로 인한 위험을 측정하고 평가하는 일도 쉽지 않다. 필요한 데이터가 무엇인지를 예상하고, 목적을 정의하고, 데이터 처리요건을 더 구체화하기 때문에 인공지능 학습에 사용되는 데이터 조합은 더 한정될 수 있다. 학습에 활용된 데이터가 한정되어 편향(bias)이 생긴다면, 학습데이터에 의해 훈련이 이루어지는 인공지능의 알고리즘을 활용한 결과도 편향될 수 있다.

우리나라를 비롯한 여러 나라의 입법례는 원래

3) 본 논문에서 인용하고 있는 유럽연합(European Union)의 General Data Protection Regulation (GDPR)과 유럽연합 회원국의 국내법에서 연구를 위한 정보보호 법률조항은 대동소이하다. 개인정보처리는 ① 적법성, 공정성, 투명성의 원칙(Lawfulness, Fairness and Transparency), ② 목적 제한의 원칙(Purpose Limitation), ③ 개인정보 최소화의 원칙(Data Minimization), ④ 정확성의 원칙(Accuracy), ⑤ 보관 기간 제한의 원칙(Storage Limitation), ⑥ 무결성, 기밀성의 원칙(Integrity and Confidentiality) 등 6대 원칙에 따라야 한다. 이에 더하여 정보 주체의 동의, 공익, 정당한 이익과 같은 적법성 요건을 명시하고 있다. 건강정보, 유전정보의 처리에 관하여는 추가적인 조건을 도입할 수도 있다(GDPR Article 9(4)).

의 목적이 달성된 후 또는 개인의 요청에 따라 개인정보를 삭제하거나 데이터 사용을 제한하게 되어 있다. 국제적으로 통용되었던 데이터 보호지침인 OECD의 '프라이버시 보호와 개인정보의 국제유통에 대한 가이드라인에 관한 이사회 권고(Guidelines on the Protection of Privacy and Transborder Flows of Personal Data)' [8]와 현대의 각국의 입법례에는 공통된 원칙과 요건이 있다. 데이터의 수집 제한(Collection Limitation Principle), 목적 명시(Purpose Specification)와 같은 원칙이나 명시적인 정보 주체의 동의 또는 해당 데이터의 익명화와 같은 데이터처리 요건은 인공지능을 통해 얻는 이점을 방해할 수 있다. 더 오랜 기간 또는 준 영구적으로 데이터를 보관하면 현재의 개인정보 보호법 제도에서 추구하는 데이터 보안에 취약해지는 반면, 인공지능 알고리즘을 최적화하려면 더 많은 양의 데이터가 필수적이다. 아래에서 개인정보보호의 원칙과 인공지능기술의 긴장 관계를 동의, 익명성, 위험 점수화 알고리즘 측면에서 살펴본다.

1. 동의와 책임에 대한 위협

데이터 처리기술이 진화하는 속도가 빨라지고, 예기치 못한 과학 지식의 연결성이 증가하는 것을 고려하면 한 가지 용도로 사용되는 것에 동의했던 데이터는 완전히 다른 목적으로 사용할 수 있다는 것을 예상할 수 있다. 여러 공공기관의 데이터베이스로부터 개인정보수집이 가능한 한국보건의료연구원과 같은 기관은 정보 주체의 직접 설명에 의한 동의를 받지 않고도 개인정보를 요청할 수 있고, 받은 데이터를 통합하는 것도 가능하다.[4] 인공지능 훈련에 이용되는 데이터는 과거의 의사결정을 반영하는 특징을 갖기 때문에 데이터의 연결성이 향상되면 인공지능 활용의 영역도 넓어진다. 전자의무기록(electronic medical record, EMR)은[5] 의료기관 간의 데이터 연결성도 향상되도록 발전하고 있다[9]. 다른 방법으로는 해석해 내기 어려운 질병의 발현이나 약물에 대한 반응에 대한 유전적 · 환경적 요인의 영향에 대해 상관관계를 밝히기 위한 코호트 조사와 같은 종단연구의 활용데이터로서 EMR에 하는 사회경제적 정보를 통합하기 위함이다[10].

문제는 여러 의료기관의 다양한 의료기기로부터 수집된 개인의 헬스 데이터를 연구자 간에 공유하는 경우, 데이터 사용 목적과 정보 주체의 동의를 조화시키기 어려울 수 있다. 정보수집 시점에서부터 인공지능에 활용될 수 있다는 광범위한 가능성을 포섭하여 그 적용 범위와 목적을 특정하여 연구에 사용한다는 설명 제시가 불가능할 수 있기 때문이다. 데이터가 처음 수집된 목적 이외에 계속되는 처리를 어디까지 허용해야 할지에 관한 문제는 정보 주체나 연구자의 이익과 관련된 논쟁이나 소유권문제로 확대될 수 있다. 그래서 특정한 목적제시가 모호할 수 있다는 것에 대한 비판적인 견해도 있고[11], 동의에 근거를 둔

4) 보건의료기술 진흥법 제26조(자료의 제공) : ① 한국보건의료연구원은 연구에 필요한 정보 수집을 위하여 국가기관 및 대통령령으로 정하는 공공기관에 대하여 자료를 제출하도록 요청할 수 있다. 이 경우 그 요청을 받은 기관은 특별한 사유가 없으면 그에 따라야 한다. ② 한국보건의료연구원은 제1항에 따라 자료를 요청할 경우 「개인정보 보호법」 제23조에 따른 민감 정보 및 같은 법 제24조에 따른 고유 식별정보 등의 개인정보가 포함된 자료의 제출을 요청할 수 있다. 이 경우 해당 국가기관 및 공공기관은 개인 식별이 가능한 부분을 삭제한 후 제출하여야 한다. 〈개정 2013.7.30.〉 ③ 제2항에도 불구하고 한국보건의료연구원은 연구를 위하여 두 개 이상의 국가기관 및 공공기관이 보유한 자료를 통합하여 분석할 필요가 있는 경우에는 국가기관 및 공공기관으로부터 개인 식별이 가능한 부분을 포함한 자료를 제출받아 자료의 통합작업을 수행할 수 있다. 이 경우 자료를 통합한 후에는 반드시 개인 식별이 가능한 부분을 삭제하여야 한다. 〈신설 2013.7.30.〉

5) 보건복지부는 '진료 정보교류사업'을 진행 중이다. 2018년 말 기준으로 이 사업에 참여하고 있는 의료기관은 상급종합병원 15개를 포함하여 2,316기관이 협력 병 · 의원으로 참여하고 있다.

데이터 활용에 대한 비판적인 평가도 있다[12]. 이러한 맥락에서 OECD는 데이터 연계를 위한 입법에는 한계가 있고, 보건의료 데이터 활용을 위해서는 거버넌스(governance) 차원의 접근이 더 중요하다고 권고한 바 있다[13].

충분한 설명에 의한 동의원칙의 기원은 Nuremberg Code로 거슬러 올라간다[14]. 정보 주체의 동의 방식은 연구 방법에 따라 변화되기를 요청받고 있다. 1996년 미국에서 연방법으로 제정된 의료보험의 이동성과 신뢰성에 관한 법률(Health Insurance Portability and Accountability Act, HIPAA)은 보호되는 의료정보(Protected Health Information)라는 개념을 도입했다. 이 정보에는 인구학적 정보, 과거 · 현재 · 미래의 물리적 또는 정신적 건강에 관한 정보, 개인에 대한 의료제공에 관한 정보, 의료제공을 위한 급여 등과 관련된 정보, 그리고 개인을 식별할 수 있다고 믿는 합리적인 근거가 있는 정보도 포함된다. HIPAA의 적용대상은 병원을 포함하여 의료행위를 제공하거나 의료행위의 결과로 대가를 받는 자 모두가 그 대상이다. 또한 의료보험회사와 결제기관에도 적용이 되며 이러한 적용 대상자에게 의료정보의 활용이나 공개와 관련한 업무의 전부 또는 일부를 제공하는 자도 적용의 대상이다. 그런데 비용 절감을 위해 보험 수혜 기관들이 자신들이 보관하던 환자들의 정보를 업무제휴자 등에 20% 이상 외주(outsourcing)를 주는 상황이 되다 보니[15], 의료정보와 관련된 자료 분석 및 데이터 통합에 관련된 자도 적용의 대상이 되었다. 데이터 활용기술의 변화와 발전에 따라 기존에는 HIPAA 준수가 요구되지 않고, 보험 수혜 기관과의 계약에 의해서만 통제되었던 업무 제휴자(business associates)나 그 하청 업체(sub contractors)에도 HIPAA가 적용되도록 한 것이다.

HIPAA는 법률의 적용대상을 확대함으로써 보호할 필요가 있는 개인정보를 맥락에 따라 구분하고, 이에 맞추어 동의 규칙도 수정하고 구체화했다. HIPAA Privacy Rule에 따라 식별 가능한 정보를 연구자에게 공개할 경우, 사전적 환자의 승인(authorization)이 필요하다. 환자 승인이란 사전 동의를 의미한다. 만약 데이터를 받는 주체가 해당 데이터의 재식별을 금지한다는 데이터 이용에 대한 합의서에 서명하면 환자의 승인 없이도 HIPAA의 수범기관은 한정 데이터 세트(limited data sets)를[6] 공개할 수 있다[16]. 즉 환자의 승인 없이도 한정 데이터 세트를 만들어서 외부에 공개할 수 있으며 그 요건은 다음 두 가지이다. 첫째, 공개 목적이 연구, 공중 보건, 또는 의료서비스 운영에 한정된다. 둘째, 해당 정보를 받는 주체와 정보를 제공하는 HIPAA의 수범 기관 사이에 데이터 이용에 대한 합의서의 서명이 있어야 한다. 헬스데이터 활용을 위한 법률 규정 중에서 과학적 연구목적이면, 조건에 따라 환자의 사전 동의가 면제되는 예외규정이라고 할 수 있다.

2013년 1월 미국 보건후생부(Department of Health & Human Services) 내 민권 담당국(Office for Civil Rights)은 개인으로부터 얻은 동의서에 추후 연구의 내용이 적절히 묘사되어 개인이 자신의 건강정보가 추후 활용될 수 있다는 점을 합리적으로 예상할 수 있는 경우에는 추후 연구를 위한 동

6) 한정 데이터 세트(limited data sets)는 HIPAA Privacy Rule에서 정의된 일정한 범위의 식별 가능한 환자 정보들이다. 일반적인 건강정보가 한정 데이터 세트가 되려면 "표면적(facial)" 식별자들이 제거되어야 한다. HIPAA Privacy Rule은 한정 데이터 세트가 되기 위해서는 다음과 같은 16가지의 식별자를 제거해야 한다고 규정하고 있다(45 CFR §§ 164.514(e)(1)).: ① 이름, ② 거리 주소(타운(town), 시, 주(state) 그리고 우편번호(zip code)는 제외), ③ 전화번호, ④ 팩스번호, ⑤ 이메일주소, ⑥ 사회보장번호(Social Sevurity Number), ⑦ 의료기록번호(medical records number), ⑧ health plan의 수혜자 번호, ⑨ 계좌번호, ⑩ 라이센스(license) 번호, ⑪ 차량(vehicle) 식별자와 일련번호, ⑫ 기기(device) 식별자와 일련번호, ⑬ URL, ⑭ IP 주소, ⑮ 생체 식별자(지문, 성문(voice print)을 포함), ⑯ 전체 얼굴이 포함된 사진 이미지 및 이에 상응하는 이미지.

의를 사전에 포괄적으로 얻을 수 있다는 견해를 밝혔다[17]. 이러한 해석은 건강정보는 미래의 특정되지 않은 연구에 전용될 수 없다는 기존태도의 변화라고 할 수 있다. 그 후 2017년 전면 개정된 Common Rule에서 고지된 동의(informed consent)의 여러 요건을 신설하였다. 동의에 필요한 설명문을 온라인에 등록하는 의무를 부과하고, 동의에 필요한 기본정보를 아홉 가지로 추가하여 연구대상자들의 이해를 높일 수 있도록 하였다[18].

문제는 일정한 조건에 따른 사전 동의 면제, 포괄적인 사전 동의나 온라인 동의 방식을 통해 활용되는 데이터는 데이터에 오류가 있거나 편향이 존재한다는 것을 나중에 알게 되더라도 이를 바로잡거나 찾아낼 수 있는 기준이 모호할 수 있다는 것이다. 사전 동의는 단순히 온라인 서비스에서 '약관'을 통해 이루어지는 서비스 계약을 의미하지 않는다. 사안마다 데이터 사용자와 정보 주체 간의 충분한 설명과 설명에 대한 이해 후에 동의가 이루어지는 것이 아닐지라도 의무사항으로서 매번 동의하는 것을 포기하기에는 너무 성급한 이유가 적어도 인공지능 활용 데이터에는 있다. 인공지능을 활용하는 이유는 과거의 데이터에 기초하여 현재를 설명하고 미래를 예측할 수 있는 특성 때문이다. 만약 왜곡된 결과를 초래하는 데이터를 여러 목적으로 활용하게 되면 인공지능의 자체의 고유한 속성인 불투명성(opacity)[19][7]으로 인해 정보 주체에게 해(harm)가 되는 결정을 발견한다고 해도 이를 당사자에게 설명하거나 검증하기 어렵게 될 수 있다.

2. 데이터 익명성의 신화

익명화된 데이터(anonymous data)는[8] 식별되거나 식별 가능한 자연인 또는 개인정보와 관련성이 없도록 더 이상 식별할 수 없게 처리한다는 개념이다. 이 개념을 통해 개인정보를 식별 가능성을 기준으로 하는 개인정보보호의 법률적 프레임워크 밖으로 끌어낸다. 익명화된 데이터는 더 이상 개인정보가 아니다. 그래서 데이터를 수집하는 단계에서부터 식별 가능성(identifiability)과 익명성(anonymity)은 중요한 의미가 있다. 연구대상자의 정체가 밝혀지지 않았거나 연구자가 용이하게 판명할 수 있지 않거나 합리적으로 사용되는 모든 수단을 통해서도 자연인을 식별할 수 없는 것으로 간주하는 것이 익명화된 데이터의 개념이다. 이론적으로 익명화 방법은 다양하지만, 데이터 처리 근거로서 익명화 방법을 구체적으로 명시한 법률이나, 익명화된 데이터가 되기 위한 정도를 다루는 별도의 규정은 찾기 힘들다. 다만 익명화 과정의 끝은 문제의 데이터 집합에서 특정 개인을 재식별하는 것이 '합리적으로 불가능'하다는 익명성을 입증하는 것이다. 인공지능 시스템의 복잡한 알고리즘에서 개인정보 처리자가[9] '개인 식별이 합리적으로 불가능한' 표준을 충족하는 데 필요한 수준은 매우 높아지리라는 것은 분명해 보인다.

의학 연구를 위해 개인의 헬스 데이터를 활용할 수 있는 방법은 정보 주체의 동의를 얻는 방법 혹은 익명화 방법 사이의 선택을 의미한다. 익명성은 합법적으로 데이터의 활용 가능성을 높여

7) 복잡도가 일정 수준을 넘어선 알고리즘의 판단과정 자체에 대한 '불가해성(inscrutability)'의 문제라고 불린다.
8) 진정으로 익명화된 데이터는 '합리적으로 사용되는 모든 수단을 통해서도 자연인을 식별 불가능'해야 하고, 익명화한 프로세스를 되돌릴 수 없어야 하고, 익명화된 데이터를 원본 데이터로 재구성할 수 없어야 한다(익명화 기법(Article 29 Data Protection Working Party, Opinion 05/2014).
9) 본 논문에서 개인정보 처리자의 정의는 업무를 목적으로 개인정보 파일을 운용하기 위하여 스스로 또는 다른 사람을 통하여 개인정보를 처리하는 공공기관, 법인, 단체와 개인 등을 말한다(개인정보 보호법 제2조(5)). 한편 유럽연합의 GDPR에서 정의하는 개인정보 처리자는 컨트롤러(controller)를 지칭하는 용어로서 구분하여 기술하였다.

주면서 다양한 출처의 데이터를 특정한 제한 없이 처리할 수 있어 연구자들이 환영할 만한 개념이라고 생각할 수 있다. 하지만 많은 형태의 의학 연구에서 익명성은 적합하지 않은 개념일 수 있다. 대부분의 의학 연구는 특정한 범위와 맥락 내에서 관련성 있는 데이터를 연결 혹은 분석 할 수 있는 준 식별자를 포함한 채 사용하기 때문이다. 인간 대상 연구와 관련되는 연구대상자(human subject)의 범위가 인체유래물(biospecimen)과 대전장유전체 시퀀싱(whole genome sequencing)으로 확대되는 것을 고려하면, 데이터 처리기술이 더 정교하게 식별가능성을 제거하더라도 익명성 보장은 점점 더 어려워질 수 있다. 유전체의 고유한 특성인 가족 공유성이 있고, 사실상 일단 공개된 유전정보는 무한히 존속할 여지가 있기 때문이다. 또한 개인 식별성을 완전히 제거하는 방법은 단순하지 않으며 심지어 실현 불가능할 수도 있기 때문에 익명성에 의존하는 데이터 보호 법률에 대한 관점은 변화가 필요하다.

가명화된 데이터(pseudonymised data)는[10] 과학적 연구목적 등의 이차 활용을 위해 GDPR에 의해 도입된 개념이다. 가명화는 이름, 사회 보장번호, 생년월일과 같은 고유한 개인 식별자를 제거하여 외형적으로 개인 식별자를 익명으로 처리된 것과 같이 정보 주체를 식별하기 어렵게 만들지만, 다른 곳에 보관된 데이터와 연결하여 목적에 따라 의도적으로 개인정보로 만들 수 있다[20]. 가명화된 데이터는 특정 개인에 속하는 것으로 식별될 가능성이 남아있으므로 따라서 개인정보이다. 가명화된 데이터의 활용이 증가할수록 개인정보처리 원칙에 근거한 데이터 최소 수집과는 어긋나는 것이다.

인공지능 시스템을 학습시키는 데이터는 과거의 데이터에 기반하며 데이터의 가치는 시간 경과에 따른 지속적인 피드백을 수집할 수 있는지에 따라 달라진다[21]. 인공지능 학습 데이터는 그 시스템 구성이 변경될 경우 반복적으로 재사용해야 하는 상황이 발생하기 때문이다. 인공지능 프로그램이 임상적으로 의미있고 일련의 환자 집단을 식별할 수 있는 실제적인 발견을 했다면, 그 발견이 다른 데이터 세트에서도 유지되어야만 더 가치 있는 데이터가 될 것이다. 이러한 특성은 재식별이 가능한 한정된 표본 데이터가 어느 정도 편향성을 가지고 있다면 그 결과는 일회에 그치지 않을 수 있다는 사실을 말해준다. 특정 치료를 권장하는 결과는 다른 특정 집단에는 건강상 위험을 초래한다거나, 단기적인 치료에는 도움이 되지만 장기간 치료를 한다면 새로운 문제가 발생할 수 있는 치료법으로 결정한다거나, 사전에 정의한 목적에는 맞지만 향후 문제를 일으키는 방향을 선택하는 경우가 생길 수 있다[22].

3. 위험점수(risk scores)와 알고리즘적 차별 (algorithmic discrimination)

피츠버그 대학교(University of Pittsburgh)는 고령자와 신체장애인를 대상으로 하는 메디케이드[11] 장기프로그램과 가정지원 서비스 여부를 결정하는 도구와 스크리닝 도구에 대한 알고리즘의 신뢰도와 유효성에 대한 연구를 수행했다[23]. 메

10) GDPR에서 도입한 가명화된 데이터(pseudonymised data)는 추가적인 정보의 사용 없이는 더는 특정 개인의 속성으로 인정되지 않는 개인정보를 의미한다. 단, 추가적인 정보는 분리하여 보관해야 하며, 개인이 식별되지 않도록 기술적 · 관리적 조치를 해야 하는 의무사항이 있다. GDPR 서문 26–29, 79 참조.

11) 미국의 공적보험인 메디케이드(Medicaid)는 저소득층 어린이와 성인, 노약자와 장애인을 포함하여 7,000만 명이 넘는 사람들을 대상으로 하는 건강 및 장기간호 프로그램이다. 연방정부와 주 정부가 공동으로 기금을 마련한다. 민간보험에 가입하지 못하는 사람들에게 보건의료 시스템의 혜택을 받게 하고, 모든 요양원 거주자의 2/3 이상을 지원하는 프로그램을 제공하여 정신건강 및 장기요양 보호를 위한 가장 큰 지불자 역할을 한다.

디케이드 청구서 데이터와 임상 데이터가 통합된 알고리즘이 청구서 데이터에만 근거한 것보다 더 정확한지 여부를 밝히는 이 연구에서 기존의 평가도구에 따라 적격판정을 받은 사람들이 알고리즘을 사용하면 자격이 없는 것으로 나타났다. 알고리즘 기반으로 적격성 기준을 효과적으로 변경할 수 있는지에 대한 고찰에서 알고리즘을 사용하여 개인의 치료 수준을 결정하는 것은 우려할 만하다는 것이 확인된 것이다[24].

호프만(Hoffman)에 따르면 알고리즘은 정확하지 않은 공개된 데이터에 의존할 수 있으며, 환자 자신의 역량은 부정될 여지가 있다. 예측 알고리즘이 결론을 끌어내는 프로그램을 사용하려면 임상의의 직감을 무시하거나 편견을 재강화하지 않는다는 보장이 필요하다[25]. 마약성 진통제 위험점수를 예로 들어보자. 마약성 진통제에 대한 문제를 해결하기 위해 예측분석을 사용하는 보험회사들이 있다. Cigna 사(社)는 과다투여 가능성이 있는 환자를 신고하는 프로그램을 확대하고 있다[26]. Optum 사(社)는 마약성 진통제 관련 위험점수로 환자를 계층화한다[27]. LexisNexis 사(社)는 마약성 진통제 사용 장애가 있는 환자 정보를 보험명세에 반영한다[28]. 만약 친척이나 동거인이 교통사고를 당하여 일시적으로 약국을 방문하면 같은 주소에 살고 있다는 자료에 근거하여 '상대적으로 강한 연결'을 도출할 수 있다. 양 당사자가 같은 의료 보험에 가입하면 소프트웨어는 '두 사람의 전체 행동'에서 하나의 패턴을 찾을 수 있다. 이러한 패턴에서 제시하는 위험점수가 마약성 진통제의 과다복용자로 인식되는 결과를 초래한다. 교통사고 환자에게 감정적 · 정신적 상태 때문에 일시적으로 마약성 진통제 사용의 권장 기간이 초과하였다는 맥락은 위험점수에 드러나지 않는다. 투약데이터는 의사가 약 처방을 할 때 정보에 입각한 결정을 하도록 돕는 것을 목표로 했지만 어떤 환자는 블랙리스트에 올라갈 수 있고 어떤 환자는 필요한 약을 투여받지 못하는 결과를 초래할 수 있다. 위험을 과소평가하거나 과대평가하면 보건의료 시스템의 여러 자원이 적절하지 않은 환자에게 집중될 우려가 있다.

알고리즘이란 작동이 일어나도록 내재하는 단계적 집합으로서 '데이터를 처리하는 규칙'이라는 방법을 통해 연산, 데이터 진행 또는 자동화된 추론이라는 과업을 수행한다[29]. 알고리즘이 내리는 자동화된 의사결정에는 우선순위 결정, 분류, 관련짓기, 필터링이라는 과정이 존재한다. 이 과정은 통계적 분석에 가깝고, 예측 불가능성이 내재하여 있다. 예를 들어 사람에 대한 예측에 사용되는 나이, 성별, 키 등과 같은 특성을 기반으로 단순히 개인을 분류하는 것은 예견된 목적과 별 상관없는 분석 결과를 보여줄 수 있다. 대체로 개인 또는 그룹에 대한 정보를 수집하고, 그 특성 또는 행동 패턴을 분석하고, 이를 특정 범주 또는 그룹으로 분류한 다음, 이를 이용하여 수행능력, 관심 사항, 예상되는 행동에 대해 예측 또는 평가하기 때문이다[30]. 통계적으로 유사한 것으로 보이는 다른 사람들의 특성들을 기반으로 하여 개인에 대한 무언가를 예측하기 위해 알고리즘 설계를 한다면 더욱 협소한 선택지와 강한 인과관계의 요소를 찾아 결정할 수 있다[31]. 하지만 위험점수를 생성하는데 사용하는 알고리즘에 정확히 무엇을 포함했는지 공개하지 않는다. 그래서 알고리즘의 투명성은 쟁점이 되고, 투명성이 확보되기 전까지 딥러닝 알고리즘을 써서는 안 된다고 주장하는 학자도 있다[32].

GDPR은 데이터 사용자에게 설명할 의무를 부여한다(GDPR recital 86, Article 24~28). 이 법률에 따르면 위험점수를 평가한 기업은 적절한 설명을 제공해야 한다(GDPR Article 22(3)). 하지만 개별적으로 자동화된 의사결정에 관한 설명을 제

공할 법적의무가 있다 하더라도 개인정보 처리자가 무엇을 설명해주어야 하는지 명확하지 않다. 알려준다고 하더라도 너무 복잡해서 이해하기가 불가능하거나 설명하기 위해 너무 요약하면 사실상 의미가 전달되지 않을 것이다. 새로운 알고리즘이 적용된다면 그때마다 그 변화를 어떻게 어떤 기준으로 정하여 설명할 것인지를 법률로 정하기도 쉽지 않은 일이다. 그래서 설명을 요구할 권리는 인공지능이 어떻게 기능하는지에 대한 설명을 들을 권리라기보다는 구체적인 결정이 왜 그렇게 내려졌는지를 이해할 권리라고 보아야 할 것이다. 어떤 데이터들을 활용했으며, 그 데이터의 출처가 어디이며, 그 의사결정 과정이 어떠하였는지 정도를 설명 받을 권리라고 할 수 있다. 하지만 이 경우에도 어떤 데이터가 활용되었는지 공개하게 되면 지금까지의 프라이버시 보호라는 개념과 상충하는 결과를 낳게 된다. 투명한 절차가 정보 주체의 권리를 보호하지 못하는 문제로 번질 위험이 있다. 따라서 이들 사이의 조화를 어떻게 달성할 수 있을지는 법규제가 아닌 다른 측면에서 해결의 실마리를 찾아야 할 것이다.

컴퓨터 과학과 법학에 전문성을 갖춘 연구자들은 알고리즘의 의사결정이 의도하지 않게 차별적인 결과를 낳게 되는 원인을 알고리즘 의사결정 단계별로 각각 찾을 수 있다고 본다[33]. 목표변수의 선택과 수집은 데이터에서부터 시작되므로 편향된 데이터 세트로 훈련된 알고리즘은 차별을 자동화할 가능성이 있다. 과거의 데이터로 학습한 채용 프로그램은 이미 존재하는 차별을 영속시키는 사례로 입증되었다[34]. 편향된 데이터 세트로 훈련된 시각 알고리즘은 여성이나 유색인종을 인식하지 못했다[35]. 이미 주어진 데이터를 완벽하게 학습하면 특정한 문제에만 최적화됨으로 앞으로 주어질 새로운 데이터를 제대로 설명하지 못하는 한계도 나타날 수 있다. 다른 알고리즘에 비하여 높은 성과를 낸 알고리즘은 다른 문제에 대해서는 상대적으로 낮은 성과를 낼 수밖에 없다는 의미가 된다[36]. 그래서 왜곡된 데이터로부터 편향되지 않은 결과가 나올 수 있는 알고리즘을 찾는 것과 편향성을 탐지하고 완화하는 방법을 찾는 것은 똑같이 중요하다[37].

2016년 세계경제포럼(World Economic Forum)[38]은 인공지능과 관련된 아홉 가지 윤리적 난제(Top 9 ethical issues in artificial intelligence)를 발표했다. 그중 하나가 인공지능은 항상 공정하지 않고, 중립적이지도 않다는 차별에 관한 문제이다. 미국 백악관(White House)이 발간한 보고서인 '빅데이터: 알고리즘 시스템, 기회와 시민권(Big data: a report on algorithmic systems, opportunity, and civil rights)' [39]에서도 알고리즘 기술이 현재의 경제적 차별을 영속화하거나 새로운 차별을 만들어 낼 수 있다는 문제를 제기한다. 기계학습 알고리즘이 차별을 발생시킬 수 있다는 인식은 중요하다. 인공지능기술 사용이 가파르게 늘어나면서 데이터 분석을 통해 실체세계를 반영하는 객관적인 증거를 찾을 수 있다는 믿음이 있지만 어떤 특정 집단은 편파적인 대우를 바꿀 기회를 잃게 될 수 있다. 어떤 특정 장소에 거주하는 사람들은 왜곡된 선입견에 저항할 수 없게 될 수 있다. 데이터에 의존하는 알고리즘적 편향성에 따라 이러한 문제는 재생산되고 더 악화할 수 있다. 더 다양하고 더 광범위하고 더 많은 양의 데이터를 수집하고 분석의 대상으로 삼으려 하는 데이터 근본주의(data fundamentalism)가[40] 생각하지 못했던 오류와 오류의 고착화가 발생할 수 있다.

Ⅲ. 데이터 윤리(Data Ethics)를 위하여

1. 데이터에서 '독' 제거하는 데이터 처리 (data processing)

데이터 윤리와 관련된 문제는 개인이 재식별되는 것에 국한되지 않고, 특정 개인이 어떠한 대우를 받게 될지를 결정하는 기준이나 사회적인 선택으로 확장될 수 있다. 정보 주체는 헬스 데이터 수집과 활용에 대해 동의하고, 주의 깊게 활용 목적을 선택하고, 때로는 동의 철회를 할 수 있는 권리가 있다. 이 권리가 EMR이 있는 병원에서, 소비자 직접 의뢰 유전자 검사에서 사용되는 게놈 시퀀싱 기계(genome sequencing machines)에서, 손목에 부착하고 있는 스마트폰 애플리케이션(smartphone applications)에서, 집안에 설치한 사물인터넷(internet of things)과 같은 다양한 기계장치에서 수집되는 데이터에 대하여 모두 같은 권리일까? 개인정보 자기 결정권을 이익과 해악을 잘 판단하면서 동시에 공익에도 의미가 있도록 행사하라고 정보 주체에게 맡겨놓는 것이 프라이버시 보호 문제를 해결하는 최선의 방법일까?

다양한 인공지능 기술 학습에 활용되는 데이터에 대하여 정보 주체의 개인정보 자기 결정권과 같은 권리를 보호하는 취지의 개인정보보호 원칙을 재검토하는 것으로는 충분하지 않다. 우리나라에서 일반인을 대상으로 하는 헬스 데이터를 어떤 연구에 사용해도 되는지, 어떻게 사용하는지에 대한 이해도를 조사한 사례는 없다. 보험회사는 헬스 데이터를 연구에 활용하기 전에 식별정보를 삭제하거나 암호화하는 등 프라이버시 보호 효과가 있는 방법을 적용한 후 공유하겠다고 보험가입자에게 고지하지 않는다. 보험회사가 위험점수를 예측하는 알고리즘을 오용하는 것에 대한 보호 장치로서 데이터 접근통제에 관한 규정을 만들고 이를 감독하는지 보험가입자는 알 수 없다.

GDPR 입법 과정에서 제29조 실무반(Working Party 29)이 내놓은 '자동화된 개인 의사결정과 프로파일링에 대한 지침(Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679)'에 의하면 인공지능 기술의 역량은 더 쉽게 프로파일을 생성하고 자동화된 결정을 내릴 수 있도록 하는 기술로서 이는 개인의 권리와 자유에 심각한 영향을 미칠 수 있다고 명시하고 있다[41]. 이러한 영향력이 발생하지 않도록 계약을 이행하거나, 합법적으로 적절한 조치를 명시하거나, 정보 주체의 명시적인 동의를 기반으로 하는 경우가 아니면, 자동화된 의사결정을 금지하고 있다. 의사결정 프로세스나 그 의사결정의 법적 효과와 관계없이 자동화된 처리의 결과가 미치는 영향력이 충분히 크거나 중요할 수 있다는 뜻이다[42]. 그리고 자동화된 의사결정이 정보 주체의 권리에 부당한 영향을 미치지 않도록 보장하고, 숨겨진 차별이나 편견의 위험으로부터 보호하기 위해 자동화된 의사결정에 굴복하지 않을 권리[12]를 도입하였다.

데이터 처리 측면으로 논의의 초점을 확대해 보자[43]. 데이터 처리의 효과는 사소하지 않다. 데이터의 수집, 보관, 제공, 공유 등 정보처리 과정마다 어떤 개인에게는 거의 영향을 미치지 않는 정보처리가 어떤 개인은 배제하고, 어떤 경우는 사생활에 머물지만, 어떤 경우는 취약계층과 같은 특정한 사회집단에 중대한 영향을 미칠 수 있다. 컨트롤러(controller)는 데이터 처리(data

12) GDPR Article 22(1): "개인정보 주체는 프로파일링 등, 본인에 관한 법적 효력을 초래하거나 이와 유사하게 본인에게 중대한 영향을 미치는 자동화된 처리에만 의존하는 결정의 적용을 받지 않을 권리를 가진다."

processing)를 수행할 때 데이터 사용에 관한 합리적인 결정을 할 수 있어야 한다. 개인정보의 이용 목적을 결정하고, 개인정보를 공개 또는 제공할지, (만약 한다면) 누구에게 공개할지 결정하고 정보 주체의 접근권 또는 다른 개인들의 권리가 적용되는지 여부와 개인정보의 보유 기간이나 업데이트 여부도 결정해야 한다. 이러한 결정을 할 때, 서로 다른 목표와 서로 다른 가치를 갖는 데이터가 존재하고, 그래서 서로 다른 이해관계에 따라 조합된 데이터 세트를 관리해야 한다는 지금까지와는 다른 새로운 인식이 필요하다. 어떤 목표에는 양면성이 있고 여기에 윤리적인 논쟁이 가능하다는 인식이다. 구체적인 상황과 이해관계에서 데이터에 근거한 불공정한 차별이 발생할 수 있다는 이해가 필요하다.

유럽위원회(European Commission)는 "과거에 증기기관이나 전기처럼 인공지능이 우리의 세계, 사회, 산업을 변화시키고 있다"라고 한다[44]. 이러한 변화는 데이터 보호를 위한 기술적 도구의 변화와 데이터 보호 평가 방법의 변화로 이어져야 한다. GDPR 제35조에 의하면, 컨트롤러(controller)는 "프로파일링을 포함하는 자동화된 처리를 기반으로 하며, 자연인과 관련된 법적인 영향을 끼치거나 유의하게 영향을 끼치는 자연인과 관련된 개인적 측면의 체계적이고 광범위한 평가"를 해야 하는 의무가 있다. 개인정보 정보처리자의 역할은 새로워 져야 한다. 단순히 허용 가능한 범위를 구분 짓는 결정이나 형식적인 표준을 만족시키는데 머무르지 않아야 한다. 개인정보 처리자는 법적 준수 문제로 개인정보 보호를 고려할 뿐 아니라 데이터처리의 시작부터 인공지능에 학습시키는 '옳은 데이터' 만드는 일에 대한 인식이 있어야 한다.

개인정보 처리자는 인공지능이 학습하는 데이터 자체를 편향되지 않게 건전하게 유지하려는 노력을 해야 한다. 이러한 의미에서 데이터를 직접 수집하지 않은 제삼자가 추구하는 정당한 이익과 데이터 주체의 이익이나 권리 사이의 균형을 테스트하는 것으로서 개인정보보호 영향평가(Privacy Impact Assessment)를 해야 한다. 개인정보보호 영향평가 과정에서 데이터를 공유하거나 이차 활용하기 전에 헌법상 보호받는 계층에 대한 영향을 검토하는 일이 될 수도 있다. 개인정보보호 영향평가 도구 또한 법률적 요건과 더불어 정보처리자의 윤리적 판단요건도 함께 제시해 주어야 할 것이다.

2. 사전적 절차로서의 법 제도와 조화

사전적(ex ante) 절차는 장래에 사람들의 태도와 인센티브가 바뀔 수 있음을 인식하고 어떤 결정으로 인해 초래될 결과를 고려하여 절차를 만드는 것이다. 더 나은 성능을 발휘하는 알고리즘에 대한 추구는 더 많은 양의 개인정보의 수집과 공유를 요구할 것이다. GDPR은 공익상의 기록보존 목적, 과학 · 역사 연구목적, 통계 목적을 위해서는 개인의 건강정보를 활용할 수 있는 정당성을 확보하였다. 정보 주체의 권리가 과학적 연구목적 달성을 '불가능하게 하거나 심각하게 훼손할 때' 민감 정보의 개인정보처리 금지에 대한 예외 규정을 두고 있다(GDPR Article 89(1)). 물론 동일 기관 혹은 다른 기관이 이차 활용 한다면, 과학적 연구(scientific research purposes)를 위한 추가 처리가 초기목적과 부합한다는 반론할 수 없는 추정이 존재해야만 한다는 조건이 붙어있지만, 원래 수집목적 이외의 목적이 양립 가능한(compatible) 경우에는 애초에 정보 수집을 허용한 법적 근거 이외의 별도의 법적 근거는 필요하지 않다(GDPR recital 50). 양립 가능성(compatibility)을 확인하기 위한 조건은 상당히 까다롭다.

수집목적과 추가 처리목적 간의 연관성, 해당 개인정보가 수집될 때의 상황, 추가 사용에 대해 정보 주체가 합리적으로 예상할 수 있는 해당 개인정보의 성격, 예정된 추가 처리가 정보 주체에게 미치는 결과, 애초 처리작업과 추가 처리작업에 적절한 안전장치 등을 고려하여야 한다(GDPR Article 6).

GDPR은 특수범주의 데이터 처리에 대한 예외 규정도 정했다. 상당한 공익상의 이유로 처리가 필요한 경우, 개인정보 보호권의 본질을 존중하고 정보 주체의 기본적 권리와 이익을 보호하기 위해 적절하고 구체적인 조치를 제공하는 경우, 예방의학 또는 직업의학의 목적으로 처리가 필요한 경우, 피고용인의 업무능력 평가 및 의학적 진단과 치료, 사회복지나 의료서비스의 제공, 유럽연합 또는 회원국 법률에 근거하거나 의료전문가와의 계약과 안전조치에 따라 치료 또는 사회복지나 의료서비스 관리를 위해 처리가 필요한 경우에는 공중보건을 위하면 특수범주의 데이터 처리 제한요건이 적용되지 않는다(GDPR Article 9(a)(g)(h)(i)). 오히려 과학적 연구, 역사연구나 통계 목적으로 처리되는 경우 정보 주체의 정보에 대한 접근권, 수정권, 처리 제한권, 이동권, 처리거부권을 제한하고, 공익을 위한 자료 보관이면 앞선 다섯 가지 권리 외에 정보 삭제권(right to erasure)이 제한된다[45].

GDPR의 법 조항에서 확인되는 것은 수집목적 내에서 활용해야 하는 원칙을 가지면서도 양립가능성과 같은 요건을 마련하여 헬스 데이터의 활용과 보호의 균형점을 모색한다는 것이다. 법적 근거인 높은 기준의 동의에만 의존하면 대표성이 없거나 편향된 데이터, 크기가 부적절한 표본데이터를 활용할 개연성이 높다. 또한 동의 철회된 데이터는 연구 결과에 영향을 줄 수도 있다. 이해관계를 정리하는 합법적인 절차로서의 법률 요건도 필요하고, 동시에 맥락과 목적이 바탕을 이루는 새로운 원칙의 등장이 필요하다.

우리나라는 개인정보 보호법에서 규정하는 민감 정보 개념에 근거하여 헬스 데이터를 활용하려다 보니 정보 주체의 동의 없는 활용은 원칙적으로 금지되어있다. 그래서 연구자들은 '생명윤리 및 안전에 관한 법률'에 근거하여 기관생명윤리위원회의 심의를 통해 연구대상자의 동의를 받는 것을 면제받고 있다(생명윤리 및 안전에 관한 법률 제16조(3)). 하지만 이 법률에 근거하는 동의 규칙은 미(未)경험 영역을 간과하고 있다는 점을 지적하고 싶다. 칼라브레시(Calabresi)와 멜라메드(Melamed) [46]는 권리침해 상황에 대한 구제책을 선택하면서 권리 유형을 구분하여 그 대응 방식을 제시했다. 이들은 권리침해에 대해 효율적으로 대처하기 위해 '권리자와 침해자 간 가치평가의 효율성'과 '권리자와 침해자 간 거래 비용의 수준'을 잣대 삼아 동의 규칙을 적용할지 보상 규칙을 적용할지를 고민했다. 그 결과, 기(旣)경험 영역에 대한 규제에 대응하기 위해서는 동의 규칙(property rule)을 적용하고 미(未)경험 영역에 대한 규제에 대응하기 위해서는 보상 규칙(liability rule)을 적용하는 방안을 제시했다. 동의규칙이란, 권리자가 권리 이용을 동의해주지 않는 이상 허용될 수 없다는 의미로서, 절대권의 성격을 지닌다. 반면 보상 규칙이란, 권리를 제한하더라도 보상만 해주면 침해가 용인된다는 의미로 상대권의 성격을 가지므로 제한 가능성이 있는 권리에 적용되는 권리 침해에 대한 구제책이다. 동의규칙은 사전 예방에 중점을 두게 되며 국가나 기타 제삼자(Trust Thirty Party)의 개입이 거의 배제된다. 하지만 보상규칙은 권리가 침해된 후 침해자에 보상책임을 지우는 방식이므로 권리가 침해될 가능성에 대비하여 보상청구권과 함께 사후적인 배제청구권이 인정된다.

미국의 경우, 헬스 데이터에 따른 다양한 형태의 차별을 방지하는 법이 있다. 장애인법(Americans with Disabilities Act), 유전정보차별금지법(Genetic Information Nondiscrimination Act), 환자 보호와 부담 적정보험법(Patient Protection and Affordable Care Act) 등이 그것이다. 형편에 따라 차별을 방지하는 법률이 마련되어있다. 하지만 누군가 소비자 직접의뢰 유전자 검사(DTC-GT)를[13] 하여 알츠하이머 유전자가 발견되었다고 가정해 보자. 유전자 검사 기관이 분석 결과를 장기보험회사와 공유하고 보험회사는 보험료를 인상한다면 이를 방지할 수 있는 법률은 없다. 미래의 연구를 위해 사전에 동의해야 하는 정보 주체에게 적용해 본다면, 동의와 동의 철회의 자유를 충분히 보장받되 만일 사고가 발생하면 그에 상응하는 책임을 부담하는 방식으로 규제가 적용되도록 하는 보상규칙을 생각해볼 수 있을 것이다. 미경험 영역은 가치평가의 효율성이 낮고 거래 비용이 크기 때문에 보상규칙이 적용되어야 효율적일 수 있다. 특히 국가나 제삼자의 개입이 전제되는 대규모 유전체 데이터를 활용하는 연구면 권리침해 자체를 허용하지 않아 권리침해를 예방하려는 동의 규칙과 권리침해 가능성은 열어두되 그에 따른 보상을 하여 권리침해를 억제하려는 보상 규칙을 적절히 도입하여 정책 방안을 마련할 필요가 있다.

3. 사후적 판단을 위한 알고리즘의 투명한 설계

인공지능 기술이 이전의 기술과 다른 특징은 투명성에 영향을 미친다는 점이다. 인공지능 시스템을 종종 블랙박스(black box)라고 일컫는 이유는 데이터를 가지고 반복된 학습을 시켜 모델을 만들 때, 복잡한 확률적 상관관계를 찾는 데이터 간의 상호 연관되는 방법이 드러나지 않기 때문이다[47]. 알고리즘의 투명성은 사람이 이를 들여다볼 수 있다는 것을 의미한다. 투명성이 있어야만 '인공지능이 조언하고 인간이 결정한다'라는 주장이 성립될 수 있다. 투명성은 오래전부터 유럽연합의 법에서 존중됐다. Treaty on European Union 제1조는 "가능한 한 시민에게 공개적이고 밀접하게 내려지는" 모든 결정에 대해 언급하고, 제11(2)조에서 "유럽연합 조직들은 시민사회나 대표 단체와 개방적이고 투명하며 정기적인 대화를 유지해야 한다"고 명시하고 있다[48]. Treaty on the Functioning of the European Union 제15조는 유럽연합의 시민들이 유럽연합 기관과 조직, 기업 등의 문서에 접근할 수 있는 권리와 그러한 유럽연합 기관과 조직, 기업 등이 정보처리를 투명하게 한다는 보장을 위해 지켜야 하는 요구사항 등에 대해 명시하고 있다[49].

투명성은 합법성뿐 아니라 공정성(fairness)을 달성하는 데 중요한 요소가 된다[50]. 공정성이 언급되는 이유는 자동화 알고리즘의 완전한 투명성을 보장하는 것이 매우 어려워 정보 주체의 고지(notice)를 받을 권리가 침해당할 수 있기 때문이다[51]. 시트론(Citron) [52]의 설명에 의하면 모바일 앱과 공유된 개인의 칼로리 섭취량 데이터에 기반을 둔 건강위험예측 알고리즘은 더 높은 보험료를 발생시킬 수 있다고 한다. 이 문제에 대한 대안으로 위험 점수화 시스템을 공개하도록 요구할 수 있는 권한을 정보 주체에게 주고 인공지능 학습 데이터가 무엇인지에 관한 내용을 일반에 공개하거나 또는 공인된 독립기관을 통해

13) DTC-GT는 소비자 직접의뢰 유전자 검사(direct-to-consumer genetic test)를 말한다. 우리나라도 '규제 샌드박스(Regulatory Sandbox)'라는 제도하여 산업통상자원부에서 수행 중인 산업융합규제 특례에서 제1호 실증 특례 대상으로 소비자 직접의뢰 유전자 검사가 지정되었다.

검증(inspect)을 받도록 하자고 제안한다. 예측 알고리즘의 공정성은 '예측 알고리즘에 대한 일정한 수준의 조사와 검토'라는 절차로써 보장될 수 있다는 것이다[53].

어떠한 데이터를 사용했는지, 통계적으로 중대한 차이가 있는지를 확인할 수 있도록 다양한 유형의 이해관계자들에게 공개하고, 필요하다면 정보 주체에게 이의를 제기할 수 있도록 하는 절차를 마련해 당사자에게 이를 알려주는 것은 중요하다. 하지만 이러한 대안도 문제는 있다. 인공지능 학습의 대상이 되는 데이터 자체를 감독하고 규제하기 위해 해당 개인정보를 공개하거나 검증을 받게 하는 방식은 오히려 정보 주체의 프라이버시를 침해할 수 있다. 현재 투명성을 위한 대안은 인공지능 시스템을 설계할 때부터 이해관계자들이 서로 협의하여 필요한 데이터를 수집하기 위한 통합된 방법을 개발하는 것이다. 헬스 데이터를 활용하는 연구 공동체를 어떻게 만들고, 데이터 수집 도구(의료기기나 기계장치)의 기준은 어떠해야 하며, 누가 어떻게 데이터에 접근할 수 있는지부터 고민해야 한다.

Ⅳ. 마치면서

헬스 데이터를 수집하는 능력 자체는 보건의료 산업 영역에서 기술혁신의 표현일 수 있다. 무엇을 먹고, 얼마나 걷고, 어떤 건강 보조기를 구입하고, 어떤 사이즈의 옷을 주문하는지와 같은 일상생활에서 수집되는 데이터는 건강의 단서가 되고 개인의 건강에 대해 많은 정보를 준다. 그렇다고 이를 모두 헬스 데이터로 포괄하여 법률로서 민감하게 보호되어야 하는 데이터로 획정한다면 긴급한 상황에서 생명을 구하기 위한 공유나 잘 알려지지 않은 특수한 상황의 사회적 격차를 제거하는 근거를 찾는 연구도 진행하기 어려울 수 있다. 고급 알고리즘은 점점 더 많이 생성될 것이고 그러한 알고리즘으로 초래될 수 있는 부작용을 예상하는 것도 수월하지 않을 것이다. 그런 의미에서 헬스 데이터 자체가 수집되는 방식을 엄격히 제한하기보다는 헬스 데이터의 활용에 관한 규제와 이를 강화하는 방안을 고민하는 방향으로 법 제도는 변화하고 있다. 하지만 모든 사람의 데이터를 보호하는 법률로 구체적인 절차에 대한 요구 사항과 이를 준수하는 방법까지 마련한다는 것은 어려운 일이다.

헬스 데이터를 활용하면서 프라이버시 침해를 막는 규제가 존재한다고 해도 모두를 만족하게 할 수는 없을 것이다. 개인마다 어떤 종류의 데이터가 민감한지 그 체감도는 다르기 때문에 누군가는 자신의 헬스 데이터가 인공지능을 더 똑똑하게 만들 수 있다면 공유되는 것을 거부하지 않을 것이고, 누군가는 사용되는 목적이 무엇인지와 상관없이 정말 사적으로 남아있길 원할 것이다. 헬스 데이터를 공유하고 분석하는 인공지능을 활용하는 방식은 법률의 경계선에서 다른 차원으로 논의를 이어가야 한다.

'데이터 윤리'와 관련된 논의는 법이나 규제가 아닌 인공지능 윤리라는 새로운 유형으로 어떠한 내용을 어떻게 담을지에 대하여 함께 고민하는 것이다[54]. 우리나라는 2008년 3월 '지능형 로봇 개발 및 보급 촉진법' 제정 후, 이 법률 제18조에 근거하여 '로봇 윤리 헌장'을 제정하려는 시도가 있었다. 로봇의 설계자, 제조자와 사용자가 주체가 되는 윤리적 패러다임을 수립하려고 했으나 공표되지 못하였다. 하지만 적어도 기술의 발전이 안고 있는 결점과 사회적 폐단에 관한 관심이 필요하다는 것에 공감하고 이를 해결하는 방식으로서 윤리적 원칙에 대한 논의를 시작한 것으로 보인다. 2018년 6월에는 '지능정보사회 윤리 가이드라인과 윤리 헌장'이 발표되었다. 데이터 윤

리에 대한 논의는 개인정보와 관련된 권리침해와 편견과 차별의 특성을 재정의하기 위한 새로운 기준을 만드는 일로 진도(進度)가 나가야 한다.

인공지능에 활용되는 헬스 데이터 윤리는 법률이나 규칙을 작성하는 사람, 데이터를 처리하고 알고리즘을 개발하는 사람, 알고리즘을 활용하여 연구하는 사람들의 의무를 고려한 윤리원칙을 만드는 일이다. 개인정보 처리자가 수행하는 개인정보보호 영향평가는 기술적 관점에서 신뢰할 만큼 견고하고 안전해야 한다. 알고리즘 개발자와 시스템 설계자는 인공지능을 작동시키는 과거의 편향된 결정으로 인해 데이터가 부정확할 수 있고, 누락되는 변수로 인해 불완전할 수 있다는 것을 이해하고 데이터 변수를 선택해야 할 것이다. 연구자들은 표본데이터를 추출하는 과정에서 데이터의 편향이 일어날 수 있다는 것을 이해하고 인공지능을 훈련하거나 테스트용 헬스 데이터 세트를 마련하는 일에 협력해야 한다. 정부나 데이터를 관리하는 기관은 데이터처리와 관리에 대한 지침을 새롭게 정할 때, 이용자에게 그 데이터를 사용하는 근거를 설명하는 요건과 그러한 설명을 어떠한 방식으로 해야 하는지 원칙을 마련해야 할 것이다.

데이터 편향을 줄이는 방법으로 인공지능이 맥락에 따라 최선의 결과를 얻기 위한 규칙을 하나씩 배우도록 하여 개별상황에서 상대적으로 올바른 선택을 하는 과정을 반복시켜서 높은 수준의 올바른 판단을 내리게 하는 방식을 생각할 수 있다. 하지만 데이터의 수집경로와 데이터에 얽힌 맥락이 다르므로 이러한 다른 속성들 사이에서 일관된 선택의 기준을 정하기 전에 역시 시간에 따라 변화된 맥락과 개인에게 미치는 영향을 찾아내고, 그 시점에서 인정되는 윤리적 원칙을 적용해야 할 것이다. 인공지능은 자연과 인공물, 지각력 없는 존재와 지각력 있는 존재의 구별을 당연시 여겼던 그동안의 사고방식에 균열을 만들고 있다. 인공지능의 '지능'을 정의하고 이것을 어디까지 발전시키느냐는 기준은 그 사회마다 다를 것이다. 이 '지능'의 목표는 기술 하나로 도달할 수 있는 것이 아니다. 절차적인 원칙을 위한 법 제도는 물론이고, 사용자와 개발자, 규제기관을 위한 윤리적 요구를 마련해야 할 것이다. 리쾨르(Ricoeur)가 호소한 '실천적 지혜(la sagesse pratique)'와[55] 같은 윤리적 목표를 가지고 상황의 특수성에 맞는 대안을 찾아가는 노력이 필요하다.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

1) OECD. Artificial intelligence in society, 2019. Available from: http://www.oecd.org/sti/ieconomy/artificial-intelligence-in-society-eedfee77-en.htm [cited 2019 Jun 17]
2) Constantiou ID, Kallinikos J. New games, new rules: big data and the changing context of strategy. Journal of Information Technology 2015 ; 30(1) : 44-57.
3) 마이클 네그네빗스키. 김용혁 역. 인공지능 개론. 서울 : 한빛아카데미, 2013 : 23-25.
4) The New York Times. Just don't call it privacy. 2018.9.23. Available from: https://www.nytimes.com/2018/09/22/sunday-review/privacy-hearing-amazon-google.html?rref=collection%2Fsectioncollection%2Fbusiness [cited 2018 Dec 23]
5) European Union (EU). 2018 reform of EU data protection rules. Available from: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en [cited 2018 Dec 23]

6) Nature. Bias detectives: the researchers striving to make algorithms fair. Available from: https://www.nature.com/articles/d41586-018-05469-3 [cited 2019 Mar 2]
7) Nature. AI can be sexist and racist — it's time to make it fair. Available from: https://www.nature.com/articles/d41586-018-05707-8 [cited 2019 Mar 2]
8) OECD. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Available from: http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm [cited 2019 Mar 2]
9) 보건복지부 보도자료. 환자 편의를 위한 진료정보교류사업 전국 네트워크 완성! 2019. 5. 30.
10) Chow-White P, Macaulay M, Charters A, et al. From the bench to the bedside in the big data age: ethics and practices of consent and privacy for clinical genomics and personalised medicine. Ethics and Information Technology 2015 ; 17(3) : 189-200.
11) European Union (EU). Data Protection Working Party Article 29, Opinion 03/2013 on purpose limitation. 2013. Available from: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf [cited 2019 Jun 29]
12) Tene O, Polonetsky J. Big Data for All: privacy and user control in the age of analytics. Nw J Tech & Intell Prop 2013 ; 11(5) : 239-273.
13) OECD. Recommendation of the Council on Health Data Governance 2016. Available from: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433 [cited 2019 Jun 17]
14) National Institutes of Health. The Nuremberg Code. Available from: https://history.nih.gov/research/downloads/nuremberg.pdf [cited 2019 Jun 17]
15) U.S. Department of Health and Human Services Office for Civil Rights. Reports to Congress on Breach Notification Program, Annual Report to Congress on Breaches of Unsecured Protected Health Information 2013-2014. Available from: https://www.hhs.gov/sites/default/files/rtc-breach-20132014.pdf [cited 2019 Mar 2]
16) Johns Hopkins Medicine. Definition of limited data set 2015. Available from: https://www.hopkinsmedicine.org/institutional_review_board/hipaa_research/limited_data_set.html [cited 2019 May 2]
17) Office of the Federal Register. Federal policy for the protection of human subjects. Federal Register 2017 ; 82(12): 7149.
18) HHS.gov. Revised Common Rule Q&As. Available from: https://www.hhs.gov/ohrp/education-and-outreach/revised-common-rule/revised-common-rule-q-and-a/index.html [cited 2019 Apr 29]
19) Selbst AD, Barocas S. The intuitive appeal of explainable machines. Fordham Law Rev 2018 ; 87 : 1094-1096.
20) Fears R, Brand H, Frackowiak R, et al. Data protection regulation and the promotion of health research: getting the balance right. Quarterly Journal of Medicine 2014 ; 107(1) : 3-5.
21) Agrawal A, Gans J, Goldfarb A. Prediction Machines: The Simple Economics of Artificial Intelligence. Boston (MA) : Harvard Business School Press, 2018.
22) Challen R, Denny J, Pitt M, et al. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019 ; 28(3) : 231-237.
23) University of Pittsburgh, Medicaid Research Center. Medicaid Managed Long-Term Services and Supports. Available from: http://www.healthpolicyinstitute.pitt.edu/medicare-medicaid/research/medicaid-managed-long-term-services-and-supports [cited 2019 Jun 29]
24) Pennsylvania Health Law Project. Updates on Community Health Choices. Health Law PA News 2018 ; 21(5) : 1-10.
25) Polico. How your health information is sold and turned into 'risk scores'. Available from: https://www.politico.com/story/2019/02/03/health-risk-scores-opioid-abuse-1139978 [cited 2019 Jun 29]
26) Cigna. Fighting the opioid epidemic. Available from: https://www.cigna.com/about-us/newsroom/studies-and-reports/fighting-opioid-epidemic [cited 2019 Jun 29]
27) Optum. Mapping patterns to fight the opioid

crisis. Available from: https://www.optum.com/health-insights/geographic-opioid-patterns.html [cited 2019 Jun 29]
28) LexisNexis Risk Solutions. LexisNexis Risk Solutions Government CEO: the role of data and analytics in helping to mitigate the opioid epidemic at HIMSS 2018. Available from: https://risk.lexisnexis.com/about-us/press-room/press-release/20180306-himss [cited 2019 Jun 29]
29) 위키백과. 알고리즘. Available from: https://ko.wikipedia.org/wiki/알고리즘 [cited 2019 Mar 2]
30) Lusignan S. Effective pseudonymisation and explicit statements of public interest to ensure the benefits of sharing health data for research, quality improvement and health service management outweigh the risks. Informatics in Primary Care 2014 ; 21(2) : 61-63.
31) Council of Europe. The protection of individuals with regard to automatic processing of personal data in the context of profiling Recommendation CM/Rec(2010)13. Available from: https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805cdd00 [cited 2019 Mar 2]
32) Open Transcripts. Making an ethical machine. Available from: http://opentranscripts.org/transcript/making-an-ethical-machine/ [cited 2019 Mar 2]
33) Barocas S, Selbst AD. Big data's disparate impact. California Law Review 2016 ; 104(3) : 671-732.
34) Reuters. Amazon scraps secret AI recruiting tool that showed bias against women. Available from: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G [cited 2019 Mar 2]
35) Gender Shades. How well do IBM, Microsoft, and Face++ AI services guess the gender of a face? Available from: http://gendershades.org/ [cited 2019 Mar 2]
36) Wolpert D, Macready WG. No Free Lunch Theorems for Optimization. IEEE Transactions on Evolutionary Computation 1997 ; 1(1) : 67-82.
37) MIT Technology Review. AI has a culturally biased world view that Google has a plan to change. Available from: https://www.technologyreview.com/the-download/612502/ai-has-a-culturally-biased-worldview-that-google-has-a-plan-to-change/ [cited 2019 Mar 2]
38) World Economic Forum. Top 9 ethical issues in artificial intelligence, World Economic Forum 2016. Available from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf [cited 2019 May 2]
39) Executive Office of the President. Big data: a report on algorithmic systems, opportunity, and Civil Rights 2016. Available from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf [cited 2019 May 2]
40) Crawford K. The hidden biases in big data. Harvard Business Review 2013. Available from: https://hbr.org/2013/04/the-hidden-biases-in-big-data [cited 2019 May 2]
41) WP 251, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Available from: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053 [cited 2019 May 2]
42) 박미정. 보건의료 빅데이터 활용에 관한 법·정책적 개선방안 연구. 한국의료법학회 2018 ; 26(1) : 165-194.
43) Kuner C, Cate FH, Svantesson DJB, et al. Machine learning with personal data: Is data protection smart enough to meet the challenge? International Data Privacy Law 2017 ; 7(1) : 1-2.
44) Communication from the Commission, Artificial intelligence for Europe. COM 2018. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN [cited 2019 Mar 2]
45) Bird & Bird. Derogations and special conditions. Available from: https://www.twobirds.com/~/media/pdfs/gdpr-pdfs/81--guide-to-the-gdpr--derogations-and-special-conditions.pdf?la=en&hash=0D16F81389A0F1D60D20BB503DF92ACC701E4571 [cited 2019 Jun 17]

46) Calabresi G, Melamed D. Property rules, liability rules and inalienability: one view of the cathedral. Harvard Law Review 1972 ; 85 : 1089.
47) Medium. Optimization over explanation – maximizing the benefits of machine learning without sacrificing its intelligence, Medium, 2018. Available from: https://medium.com/@dweinberger/optimization–over–explanation–maximizing–the–benefits–we–want–from–machine–learning–without–347ccd9f3a66 [cited 2019 Jun 17]
48) Official Journal of the European Union. Consolidated version of the treaty on european union. 2012. 10. 26.
49) Wikisource. Consolidated version of the Treaty on the Functioning of the European Union 2012. Available from: https://en.wikisource.org/wiki/Consolidated_version_of_the_Treaty_on_the_Functioning_of_the_European_Union/Part_One:_Principles#Article_15 [cited 2019 Mar 2]
50) What–If Tool. Playing with AI fairness: Google's new machine learning diagnostic tool lets users try on five different types of fairness. Available from: https://pair–code.github.io/what–if–tool/ai–fairness.html [cited 2019 Jun 17]
51) Crawford K, Shultz J. Big data and due process: toward a framework to redress predictive privacy harms. Boston Law Review 2014 ; 55 : 101–125.
52) Citron DK. Technological due process. Washington University Law Review 2008 ; 85 : 1249–1313.
53) Citron DK, Pasquale F. The Scored Society : due process for automated predictions. Washington Law Review 2014 ; 89(1) : 15–16.
54) The New York Times. Is ethical A.I. even possible? Available from: https://www.nytimes.com/2019/03/01/business/ethics–artificial–intelligence.html [cited 2019 Mar 2]
55) Ricoeur P. [Soi–même comme un autre]. Paris : Éditions du Seuil, 1990 : 279–290. French.

한국의료윤리학회지 제22권 제3호(통권 제60호) : 255–273 ⓒ한국의료윤리학회, 2019년 9월
Korean J Med Ethics 22(3) : 255-273 ⓒ The Korean Society for Medical Ethics, September 2019
pISSN 2005-8284 eISSN 2234-3598

A Study on Artificial Intelligence and Data Ethics: Focusing on Health Data Used in Artificial Intelligence

PARK Mi Jeong*

Abstract

Health data collected in various ways and forms is secondary use of scientific research purposes. Including Republic of Korea, several country's Data protection law require the anonymity of data in addition to obtaining the consent of the data subject as a provisions relating to specific data processing situations. In the case of research related to the important objectives of public interest, the informed consent of the subject shall be exempt from that liability. In order to find such compatibility of purposes, consideration will need to be taken in terms other than the lawful processing of personal data. This paper starts with the fact that data is used to train artificial intelligence. First, artificial intelligence needs to focus on specifics on what data are used in the training of the artificial intelligence and how the algorithms are built, and the concerns arising from the mechanism of algorithms are discussed. The data used in the artificial intelligence system are considered as the subject of ethical debate and the ethical problems are discussed. And analyzed the legislative process and legal provisions for data processing principles of EU General Data Protection Regulation to find clues to solving that problems. The problems that can arise due to the characteristics of artificial intelligence technology, which are hard to prepare and interpreter through legislation, are explains to informed consent and responsibility, the myth of data anonymity and, risk scores and algorithmic discrimination. As a conclusion, I suggested about data processing to removes 'poison' from the data, harmonization with legal system as an ex ante procedure, and transparent design of the algorithm for human judgment as a whole.

Keywords

artificial intelligence, algorithms, anonyms and pseudonyms, ethics, health

* *Senior Research Fellow, Center for Education on Healthy Society, College of Medicine, Seoul National University:* ***Corresponding Author***