

데이터마이닝(TDM)과 특허권의 효력제한

권지현*

- I. 서론
- II. 데이터특허의 개념과 유형
 1. 데이터특허의 개념
 2. 데이터특허의 유형
- III. 데이터마이닝과 특허권의 효력 문제
 1. 데이터특허의 효력과 보호범위
 2. 데이터마이닝과 특허권의 효력제한
 3. 주요국 특허법상 특허권의 효력제한
- IV. 데이터마이닝을 통한 데이터의 이용 극대화 방안
 1. 시험·연구에 데이터마이닝을 포함하는 방안
 2. 생산방법 추정된 물건에 데이터를 포함하는 방안
 3. 데이터의 물건 인정 방안
- V. 결론

* 광운대학교 법학부 조교수

초록

데이터 생성·수집방법특허발명 등(데이터특허권)의 실시에 의하여 얻어진 데이터는 정보재로써 상거래의 대상이 됨과 동시에 데이터구조로 가공되어 거래가 이루어지고 있다. 여기서 제3자가 특허권자의 허락 없이 특허 데이터의 속성이나 특성을 분석하는 데이터마이닝은 당해 특허권을 침해하는 것인지가 쟁점이 될 수 있다.

특허법 제96조에는 시험 또는 연구를 위하여 특허발명을 실시하는 경우 특허권의 효력이 제한된다고 규정하고 있지만, 데이터마이닝이 시험 또는 연구에 해당하는 것인지 명확하지 않으므로 데이터 활용에 의한 AI 적용 신산업이 보다 적극적으로 추진되지 못하는 문제점이 발생하게 된다.

이러한 문제점을 해결하기 위한 방안으로는 데이터마이닝도 시험 또는 연구에 포함되도록 하는 개정 방안의 검토가 필요하다. 이와 더불어 데이터특허권자의 보호강화 차원에서는 생성·수집방법특허의 실시에 의하여 생성된 데이터를 물건 또는 매체에 저장된 물건으로 간주하여 생산방법의 추정 규정이 적용될 수 있도록 제129조도 함께 개정할 필요가 있다. 또한 물건발명의 실시범위에 해당하는 물건에 데이터가 포함되도록 제2조 제3호를 개정하는 방안도 추가로 검토할 필요가 있다.

주제어

데이터, 데이터생성, 데이터수집, 데이터구조, 시험 또는 연구, 특허권의 효력 및 효력제한

I. 서론

미국 기업 오픈AI의 챗GPT 및 달리(DALL-E2), 딥마인드의 알파폴드(AlphaFold), 구글의 바드(Bard) 등과 같이 대화형·생성형 인공지능(Artificial Intelligence, AI)에는 선행 데이터(data)가 전제되어야 한다. 데이터는 AI 관련 기술의 적용과 응용에 필수적으로 요구되는 매우 중요한 정보재(information goods)이다. 산업에 석유와 전기가 필요하고 제품에 반도체가 필요하듯이, AI 기술을 적용한 산업·제조·유통·금융·의료·환경 등의 비즈니스 혁신에는 데이터가 꼭 필요하다. 데이터는 가공하지 않은 그 자체로도 상거래의 대상이 될 수 있어서 경제적 이익을 얻을 수도 있지만, 데이터 자체보다는 데이터가 가지고 있는 통계학적 특성이나 속성 등을 가공·분석하여 실제 AI학습에 적용함으로써 경제적 고부가가치를 얻을 수 있다.¹⁾

데이터는 그동안 일상생활이나 산업현장에서 자연적으로 발생하는 것을 수집·정리한 통계적 수치나 의미로서의 정보 자료에 해당했지만, 오늘날에는 자연적 발생의 데이터뿐만 아니라 특정한 목적으로 활용하기 위하여 계획적·의도적으로 카메라나 센서, 인터넷이나 네트워크 등을 통해 데이터를 수집하는 경우가 많다. 수집된 데이터는 인공지능망 학습용 데이터구조의 가공에 기본이 되고, 그 데이터구조를 바탕으로 학습하여 AI기반 소프트웨어(SW, 학습완료모델)가 창출된다. 따라서 데이터는 AI기반 SW 및 그 응용에 필수적으로 요구되는 귀중한 정보재이다.

이처럼 데이터는 미래의 먹거리 신산업 추진 및 신제품 개발에 중요한 정보재이기 때문에 (i) 데이터의 생성·수집방법 및 수집장치(시스템), (ii) 수집된 대용량 데이터 및 빅데이터(big data)의 전송방법 및 전송장치, (iii) 대

1) 2022년 데이터산업 시장 규모는 전년 대비 18.7% 성장한 20조 24억 원이다. 2021년에는 23조 972억 원, 2019년부터 2021년까지의 3년간 연평균 성장률(Compound Annual Growth Rate, CAGR)은 17.1%로 나타나면서 지속적인 성장세를 이어갈 것으로 조사되었다(한국데이터산업진흥원, “2022 데이터산업백서”, 통권 제25호, 한국데이터산업진흥원, 2022, 88면).

용량 데이터의 저장·관리방법 및 그 장치, (iv) 데이터를 가공한 데이터구조 등에 관한 기술개발이 활발히 이루어지고 있고, 이러한 데이터 관련 기술개발의 성과물은 특허출원 및 특허등록(이하, “데이터특허”) 되고 있다.²⁾

특히 데이터특허의 보호범위(제97조)와 데이터특허의 효력(제94조)에 대해서는 데이터를 물건으로 보지 않는 것으로 해석하여 누구나 데이터를 자유롭게 이용할 수 있도록 하는 것이 좋은지 아니면 “프로그램을 특정 요건³⁾하에서 물건으로 간주하는 것”과 같이 데이터의 개념을 물건 또는 방법으로 특정하여 데이터의 보호범위와 효력을 확장하는 것이 좋은지가 쟁점이 되고 있다. 데이터 수집방법의 특허발명을 업으로 실시⁴⁾하는 행위가 특허침해가 된다는 해석론은 일반적인 침해법리이고, 만약 데이터 생성방법발명의 특허발명의 실시(방법의 사용)에 의하여 생성된 데이터를 물건으로 보는 경우에는 그 데이터수집방법특허의 실시에 의하여 생산된 물건에 해당하는 데이터에도 특허권의 효력이 미치는 것으로 추정⁵⁾되기 때문에 데이터가 물건 또는

2) 발명의 명칭 및 청구범위에 “데이터, 데이터구조, 데이터세트, 데이터셋”이 기재된 특허출원 및 특허등록의 건수를 검색하면 다음과 같다[특허정보검색서비스, <http://kportal.kipris.or.kr/kportal/search/total_search.do>, 검색일: 2023. 5. 2.]

구분(10년간)(2013. 1. 1.-2022. 12. 31.)	특허출원 건수(건)	특허등록 건수(건)
데이터	발명의 명칭	26,479
	청구범위	19,888
데이터구조	발명의 명칭	232,900
	청구범위	174,880
데이터세트	발명의 명칭	375
	청구범위	351
데이터셋	발명의 명칭	21,955
	청구범위	16,835
데이터셋	발명의 명칭	105
	청구범위	81
데이터셋	발명의 명칭	18,266
	청구범위	13,207
데이터셋	발명의 명칭	145
	청구범위	96
데이터셋	발명의 명칭	2,384
	청구범위	1,578

3) 청구항의 말미에 기재된 프로그램이 물건발명으로 특정되기 위한 요건은 “하드웨어(컴퓨터 등)와 결합하여 특정 과제를 해결하기 위하여 매체에 저장된 프로그램”이어야 한다 [특허청, “기타(구 기술분야별 심사기준)-제10장 컴퓨터 관련 발명”(이하 “컴퓨터발명심사기준”), 특허청(편), 기술분야별 심사실무가이드, 특허청, 2023, 49면].

4) 여기서 특허발명의 실시란 특허법 제2조(정의) 제3호의 “가목 내지 다목”에서 규정한 것을 말한다.

5) 특허법 제129조에는 ‘물건의 생산 발명’에 대한 추정 규정이 있다(“물건을 생산하는 발

매체에 저장된 물건으로 취급할 수 있는 것인지에 따라 특허권의 효력범위도 달라진다.

데이터와 유사한 개념으로 “프로그램”의 특허보호는 프로그램방법특허 이외에도 ‘매체에 저장된 물건’ 또는 ‘하드웨어와 결합한 프로그램’의 청구항으로 기재한 경우, 그 “프로그램”을 물건발명으로 간주하여 특허를 받을 수 있다.

이러한 프로그램의 특허보호 방안을 데이터의 특허보호에 적용하여 보면, 데이터를 ‘매체에 저장된 물건’ 청구항 또는 ‘하드웨어와 결합한 데이터’ 청구항으로 청구범위를 기재한 경우에는 그 데이터를 물건발명으로 특허를 받을 수 있는지가 쟁점이 될 수 있다.

만약 데이터를 프로그램과 같이 특정 요건하에서 물건으로 간주하는 경우, 특허권자의 허락 없이 “데이터의 수집방법” 특허의 실시로 데이터를 수집하는 제3자의 행위는 특허침해가 되고, 그뿐만 아니라 수집된 데이터에도 특허권의 효력이 미치는 것으로 추정되어 특허침해가 성립하게 된다.⁶⁾ 따라서 제3자는 데이터의 속성·특성 등을 해석하는 데이터정보분석, 즉 데이터마이닝(Text and Data Mining, TDM, 이하 “데이터마이닝”)은 데이터방법특허에 의하여 생산된 데이터(물건)의 사용에 해당하여 특허침해가 되는지, 아니면 특허법 제96조 제1항 제1호에서 규정한 ‘연구 또는 시험’에 해당하여 특허권의 효력이 제한되는 것인지가 명확하지 않다. 이는 데이터를 자유롭게 이용할 수 없으므로 AI 적용 데이터산업발전을 적극적으로 촉진할 수 없고 오히려 저해하게 되는 문제가 있게 된다.

이 글에서는 이러한 문제를 해결하기 위한 방안으로 우선 데이터특허의 개념과 유형, AI발명에 있어서 데이터의 의의와 역할을 검토하고, 데이터특허권의 효력범위와 그 효력범위가 제한되는 “시험 또는 연구”와 데이터마이닝의 관계를 외국 제도와 함께 검토한다. 이러한 검토를 바탕으로 데이터특

법의 발명에 관하여 특허가 된 경우에 그 물건과 동일한 물건은 그 특허된 방법에 의하여 생산된 것으로 추정한다”).

6) 특허법 제129조의 생산방법의 추정 규정에 의하여 당해 데이터(물건으로 보는 경우)가 특허출원 시점에 국내에서 존재하지 않으면 특허침해가 성립하게 된다.

허권자와 데이터이용자의 형평성 보호 관점에서 데이터의 이용을 극대화할 수 있는 특허법 개정 방안을 제시하여 본다.

II. 데이터특허의 개념과 유형

1. 데이터특허의 개념

데이터의 생성방법·장치 및 수집방법·장치 등에 의하여 생성·수집된 데이터 또는 빅데이터는 통상적으로 Missing Data이거나 Outlier Detection이 포함되어 있으므로, 데이터 그 자체만으로는 AI학습을 수행할 수 없다. 인공지능망기술 기반 딥러닝(deep learning)⁷⁾ 학습(이하 “AI학습”)으로 사용하기 위해서는 (i) 데이터에 내재되어 있는 데이터의 규격 또는 표준화, 벡터화, 정규화, 증분 등의 속성이나 특성을 분석하는 데이터마이닝이 있어야 하고, (ii) 그다음으로 데이터마이닝의 결과를 근거로 특정 목적이나 과제수행(AI학습 목적)을 할 수 있도록 데이터구조(data structure)⁸⁾ 또는 데이터세트(data set)⁹⁾를 가공하여야 하며,¹⁰⁾ (iii) 이러한 데이터구조를 근거

7) 딥러닝 학습을 위해서는 데이터가 있어야 하고, 그 데이터의 논리적 구조인 데이터구조(데이터세트, 셋)를 가공하여야 한다. 데이터구조는 테스트의 세트(set)를 기본으로 조정이 가능한 파라미터를 사용하여 특정 과제의 해결을 위한 적합 목적함수(objective function)를 규정하고(의사결정 규칙 특징), 복수의 파라미터에서 최적 해결 방법에 접근할 수 있는 학습 알고리즘을 근거로 테스트구조를 바탕으로 학습하여 최적 ‘학습완료모델(trained model)’을 도출한다[권지현, “비기술적 요소가 포함된 AI발명의 진보성 판단기준”, 『법학논총』, 제48집(2020), 43면].

8) 데이터구조란 데이터의 속성이나 특성을 분석하고 그 분석 결과에 근거하여 AI학습용으로 가공한 것으로 “데이터 요소 사이에 각 상호관계의 속성이나 특성을 표시한 데이터 논리적인 구조(리스트, 연산, 배열, 그래프, 트리, 테이블, 큐, 스택, 파일)를 말한다[특허청, 앞의 책(컴퓨터발명심사기준), 43-44면].

9) 특허출원에서는 청구항의 말미 또는 구성요소로 “데이터구조 또는 데이터세트(셋)”란 용어를 사용하고 있으며, 특히 데이터세트(셋)는 데이터의 집합물로 정의하고 있다[곽충목·차상욱, “인공지능(AI)관련 발명의 지식재산권법상 보호방안: 특허법 및 영업비밀보

로 AI 기계학습(딥러닝)을 통하여 최적의 학습완료모델(AI기반 SW, 비즈니스 모델)이 생성되고, (iv) 사용자 또는 이용자가 특정 과제 해결에 필요한 입력 데이터를 학습완료모델에 입력하면 AI 생성물이 생성되는 것이다.

이러한 데이터와 관련하여 현행 특허법상 특허를 받을 수 있는 발명은 (i) 물건(장치, 물질)발명, (ii) 방법발명, (iii) 물건을 생산하는 방법발명이 있고, 이를 구분하기 위해서는 청구범위의 청구항 말미에 물건(장치, 물질) 또는 방법인지를 명확히 특정할 수 있도록 발명의 카테고리를 기재해야 한다. 예를 들면, 특허출원서의 청구항 말미에 「...데이터」라고는 기재할 수 없다. 이는 데이터 그 자체가 물건에 해당하는 것인지 방법에 해당하는 것인지를 특정할 수 없기 때문이다. 따라서 특허법 제42조 제4항 제2호의 규정에 의하여 특허를 받을 수 없다. 다만, 현행 특허심사기준¹¹⁾에 의하면 청구항 말미에 「...데이터구조가 저장된 기록매체」¹²⁾라고 기재한 경우에는 물건발명으로 인정하고 있으며, 또한 특허심사기준에 명문화되어 있지 않지만, 특허심사 실무에서는 「...데이터구조」¹³⁾ 또는 「...데이터세트」¹⁴⁾라고 기재한 경우에도 물건발명으로 특허를 받을 수 있다.¹⁵⁾

이처럼 데이터는 물건인지 방법인지를 특정할 수 없으므로 발명의 대상(카

- 호법을 중심으로, 『Issue Paper』, 제2019-11호(2019), 5면; 이규호, “인공지능 학습용 데이터세트 보호를 위한 특허법상 주요 쟁점 연구”, 『산업재산권』, 제64호(2020), 98면.
- 10) 일본 정부의 정책보고서에는 데이터세트(셋)를 훈련데이터세트(셋)라고 정의하고 있다 [經濟産業省 情報經濟課, 『AI・データの利用に関する契約ガイドラインと解説』, 商事法務, 2018, 263頁].
- 11) 특허청, “제1부 인공지능분야”(이하 “AI발명심사기준”), 특허청(편), 기술분야별 심사실무가이드, 특허청, 2023, 1206면.
- 12) 예로서, 청구항을 「A 구조와 B 구조와 C 구조…로 이루어진 데이터를 기록하여 하드웨어(컴퓨터)로 읽을 수 있는 매체」라고 기재한 경우.
- 13) 예로서, 청구항을 「A 구조, B 구조, C 구조…라는 데이터의 논리적 구조로 이루어진 데이터구조」라고 기재한 경우.
- 14) 예로서, 청구항을 「A 구조, B 구조, C 구조…라는 데이터의 논리적 집합물로 이루어진 데이터세트」라고 기재한 경우.
- 15) 특허등록 10-2003318호(등록일: 2019. 7. 18.) 제8항에는 “데이터구조”의 카테고리를 물건발명으로 인정하고 있고, 특허등록 10-2050755호(등록일: 2019. 11. 26.) 제11항에는 “데이터세트”를 말미에 기재하여 물건발명으로 인정하고 있다.

테고리)은 아니지만, 그 데이터의 생성, 수집, 전송, 관리 및 저장 등에 관한 방법 및 장치는 발명에 해당하여 특허를 받을 수 있다.

2. 데이터특허의 유형

(1) 데이터의 생성방법 및 수집방법

데이터에는 산업·사무 현장이나 생활환경 등에서 항시 생성되어 수집되는 정형데이터,¹⁶⁾ 센서·카메라 또는 인터넷·웹사이트 등을 통해 수시로 수집되는 비정형데이터¹⁷⁾¹⁸⁾가 있으며,¹⁹⁾ 이러한 데이터의 생성방법 및 그 장치,²⁰⁾ 데이터의 수집방법 및 그 장치²¹⁾는 특허를 받을 수 있다. 물론 단순히

16) 정형데이터란 생산·노무·부품 공급·재고·고장·수리 등 통상의 산업업무, 고객과 거래처, 매출과 회계, 진료와 보험, 성적 등과 같이 통상의 업무에 생성되는 정형적인 데이터를 말한다.

17) 비정형데이터란 센스나 디지털카메라 등에 의하여 수시로 수집되는 교통 또는 날씨·기후 등의 데이터, 수시로 발생하는 생산·작업관리·고장 등의 데이터, GPS의 위치정보데이터, 웹페이지나 SNS·블로그 등에서 수시 크롤링해서 수집되는 정보데이터, 매체에 의한 수시의 뉴스, 게시판이나 메시지 등에서 수시로 수집되는 데이터, 동영상 데이터나 음원 데이터 등과 같이 사용자가 수시 입력에 의하여 발생하는 데이터를 말한다.

18) 인터넷이나 SNS 등의 웹페이지 내용을 그대로 가지고 와서 특정 사용 목적에 따른 데이터를 추출하여 구축하는 데이터수집방법을 크롤링(crawling)이나 스크레이핑(scraping)이라고 하고, 이러한 크롤링은 데이터의 수집방법에 관한 것으로 특허를 받을 수 있다(예로서, 특허등록 10-1954813호, 등록일: 2019. 2. 27., 발명의 명칭: 하이브리드 크롤링을 통해 데이터를 수집하는 방법, 장치 및 그를 이용한 서버).

19) 일반적으로 데이터는 (1) 생산·제조·관리, (2) 유통·판매, (3) 건강·보건·의료, (4) 교통·이동·위치, (5) 은행·증권·보험·투자·자산, (6) 정보통신·미디어, (7) 도시·주택·공간, (8) 환경·기후·에너지, (9) 연구·분석·논문, (10) 문화·관광·지역 특성 등 10개의 유형으로 분류하고 있다(권태복, 『제4차 산업혁명과 특허전략』, 한국지식재산연구원·세창출판사, 2019, 232-233면).

20) ‘데이터의 생성방법’에 관한 특허권의 예는 다음과 같다.

(1) 특허등록번호 10-2429809(등록일: 2022. 8. 2., 발명의 명칭: 지형의 분석을 위한 지형 데이터 생성 방법 및 지형 데이터의 생성 방법을 수행하는 장치)

(2) 특허등록번호 10-2476446(등록일: 2022. 12. 7., 발명의 명칭: 미술 심리분석용 데이터를 생성하고, 사용자의 심리상태와 관련된 정보를 제공하는 방법)

21) ‘데이터의 수집방법’에 관한 특허권의 예는 다음과 같다.

또는 통상적으로 누구에 의해서나 동일한 데이터가 생성·수집되는 경우에는 자연법칙을 이용한 기술적 사상에 해당하지 아니하여 특허를 받을 수 없다. 데이터의 생성 및 수집방법이 특허를 받기 위해서는 데이터의 특성과 그 데이터의 생성이 특정 수단에 의하여 이루어지고 있는 기술 사상, 그 생성된 데이터의 수집이 단계별로 SW에 의한 정보처리가 되고, 이러한 정보처리는 하드웨어를 이용하여 구체적으로 실현되어야 하며, 그 정보처리의 각 단계에 대해서는 블록도 및 흐름도(플로차트)에 의하여 구현되는 기술 사상을 청구범위에 특정하여야 한다.

(2) 데이터의 전송방법 및 저장·관리방법

오늘날 데이터는 센서에 의한 수시 정보 수집, 디지털카메라 또는 스마트폰에 의한 다양한 데이터의 촬영, 멀티미디어 콘텐츠의 폭발적 증가, 인터넷 또는 소셜미디어의 활성화 등으로 대용량의 데이터(빅데이터; big data)를 실시간 수집하고, 그 수집된 빅데이터는 특정한 저장·관리센터 등에 의해 실시간으로 전송된다. 그리고 실시간으로 전송된 빅데이터를 저장·관리하는 방법 및 장치(시스템)가 있어야 한다. 이러한 빅데이터의 전송방법 및 그 장치,²²⁾ 빅데이터의 저장·관리방법 및 그 장치²³⁾ 등에 관한 기술은 특허 대상

(1) 국제출원 PCT/US2005/039489(권리자: 구글, 발명의 명칭: 크롤링 문서 및 문서 액세스 통계량 같은 데이터…광고 호스트 결정)

(2) 특허등록번호 10-2476446(등록일: 2022. 12. 7., 발명의 명칭: 미술 심리분석용 데이터를 생성하고, 사용자의 심리상태와 관련된 정보를 제공하는 방법)

22) ‘데이터의 전송방법’에 관한 특허권의 예는 다음과 같다.

(1) 특허등록번호 10-2489203(등록일: 2023. 1. 12., 발명의 명칭: 사물 인터넷 시스템에서 데이터를 전송하는 방법 및 장치, 게이트웨이 디바이스와 이의 저장)

(2) 특허등록번호 10-2471298(등록일: 2022. 11. 23., 발명의 명칭: 데이터 전송 방법, 데이터의 사용 제어 방법 및 암호 장치)

23) ‘데이터의 저장·관리방법’에 관한 특허권의 예는 다음과 같다.

(1) 특허등록번호 10-2489203(등록일: 2023. 1. 12., 발명의 명칭: 사물 인터넷 시스템에서 데이터를 전송하는 방법 및 장치, 게이트웨이 디바이스와 이의 저장)

(2) 특허등록번호 10-2471298(등록일: 2022. 11. 23., 발명의 명칭: 데이터 전송 방법, 데이터의 사용 제어 방법 및 암호 장치)

이 되고 있다. 예를 들면, 자율주행차에 필요한 강남사거리의 교통데이터를 센서나 카메라로 수집한 대용량의 데이터를 저장관리센터로 수시로 전송하기 위해서는 고도의 전송기술이 필요하고, 이러한 데이터의 전송기술은 특허 대상이 된다.

(3) 데이터의 분석·가공방법과 가공한 데이터구조(세트)

빅데이터를 AI학습 등에 활용하기 위해서는 빅데이터의 속성에 관한 분석 및 그 분석을 통한 논리적 데이터구조의 가공이 전제되어야 한다. 이와 관련한 기술로서는 빅데이터 처리 인프라를 기반으로 하는 빅데이터의 분석·가공 플랫폼에 관한 것으로 총칭되며, 특히 수집된 빅데이터의 통합기술,²⁴⁾ 빅데이터의 분석기술,²⁵⁾ 빅데이터분석의 가시화기술,²⁶⁾ AI학습을 위한 데이터의 전처리기술²⁷⁾ 등은 특허 대상이 된다. 또한 수집된 데이터를 학습할 수 있도록 가공한 데이터구조(세트)²⁸⁾도 청구범위 기재요건(제42조 제4항 제2호)을 충족하는 경우에는 발명으로써 특허 대상이 된다. 결국 이들 기술의 뒷받침은 데이터구조와 AI학습용 프로그램(딥러닝)을 통하여 생성되는 AI기반 SW기술(AI학습완료모델)이라고 할 수 있다.²⁹⁾

24) 빅데이터의 통합기술이란 다양한 방법으로 수집한 대용량의 데이터의 소스와 속성에 따라 분류하는 통합기술로서, 특히 분산된 정형데이터 또는 실시간 생성되는 비정형데이터를 수집하고, 그 수집된 대용량 데이터가 가지고 있는 속성이나 소스를 특정 틀이나 프로그래밍에 의해 자동으로 통합하는 기술을 말한다.

25) 빅데이터의 분석기술이란 통계학이나 전산학 관점에 대해서는 별론으로 하고 AI학습용 데이터구조의 가공을 위하여 데이터가 가지고 있는 속성이나 특성(규격화, 벡터화, 정규화, 표준화)을 분석하는 데이터마이닝에 관한 것을 말한다(예, 특허등록번호 10-2416558, 등록일: 2022. 6. 29., 발명의 명칭: 영상 데이터 처리방법, 장치 및 판독 가능 저장매체).

26) 빅데이터분석의 가시화기술은 데이터의 속성이나 특성을 분석(데이터마이닝)하기 위한 환경을 제공하는 데 필요한 분석 도구 기술, 그리고 그 분석 결과를 통계적으로 표시한 직관적인 정보를 관리·제공하는 인포그래픽스 기술 등을 말한다.

27) 데이터 전처리기술은 데이터의 분석기술과 AI학습으로 가공된 데이터구조를 말한다.

28) 데이터구조란 “데이터 요소 사이의 상호관계를 표시한 논리적 구조”를 말하고, 이러한 데이터구조의 가공하는 기술을 데이터구조의 가공기술이라고 한다(특허청, 앞의 책(컴퓨터발명심사기준), 43-44면).

III. 데이터마이닝과 특허권의 효력 문제

1. 데이터특허의 효력과 보호범위

데이터특허는 데이터의 생성 및 수집, 전송, 저장, 관리, 그리고 데이터의 분석, 데이터구조 또는 데이터세트 등의 기술적 사상(이하, “아이디어”)에 관한 것이다. 데이터 관련 아이디어를 특허받기 위해서는 청구항의 말미에 기재된 발명의 카테고리로서 물건(또는 장치, 시스템도 포함) 또는 방법, 물건을 생산하는 방법으로 특정하여야 한다. 이와 같은 데이터특허권의 효력은 데이터 관련 발명을 물건발명으로 보는 경우와 방법발명(물건을 생산하는 방법발명 포함)으로 보는 경우로 구분하여 생각할 수 있다.

첫째, 데이터특허를 방법발명으로 보는 경우이다. 여기서 방법발명으로서의 데이터특허란 데이터의 생성방법, 수집방법, 전송방법, 저장방법, 관리방법, 그리고 데이터의 분석방법 등을 말한다. 방법발명의 특허권 효력은 업으로 “그 방법의 사용”을 할 수 있는 권리를 독점하고(제94조 제1항), 그 방법이 특허발명이라는 것을 인지하면서 그 방법을 사용하기 위하여 청약하는 행위에도 미치게 된다(동조 제2항). 이외에도 “물건을 생산하는 방법특허”의 경우에는 그 방법에 의하여 생산된 물건에도 특허권의 효력이 미치게 된다(제129조). 따라서 데이터를 물건으로 보는 경우에는 데이터의 생성방법 및 수집방법, 전송방법, 저장방법, 관리방법, 분석방법이라는 특허발명을 사용하여 얻어진 데이터에도 특허권의 효력이 미치게 된다.

둘째, 데이터를 물건으로 보는 경우이다. 물건발명의 데이터특허란 데이터의 생성장치, 수집장치, 전송장치, 저장장치, 관리장치, 그리고 데이터의 분석장치(이하 “데이터생성장치 또는 데이터수집장치”) 등을 말한다. 이외에도 데이터구조 또는 데이터세트, 데이터구조 매체를 카테고리로서 청구항의 말미

29) 권지현, “AI발명의 성립성과 진보성 판단의 조화방안”, 『홍익법학』, 제22권 제3호(2021), 574면.

에 특정한 경우에도 물건발명으로 취급하고 있다.³⁰⁾ 물건발명의 특허권의 효력은 특허를 받은 물건, 장치(시스템)를 업으로 실시할 수 있는 권리를 독점하는 것으로, 데이터생성장치 또는 데이터수집장치, 그리고 데이터구조 또는 데이터세트, 데이터구조 매체를 업으로 실시할 수 있는 것을 말한다. 따라서 데이터특허권자의 허락 없이 업으로 데이터생성장치 또는 데이터수집장치를 생산, 사용, 양도하는 행위는 특허침해가 되고, 그리고 침해된 데이터의 생성장치 또는 수집장치에 의하여 생성된 데이터의 사용 및 양도 등의 행위에도 특허권의 효력이 미치게 된다. 또한 데이터생성장치 또는 데이터수집장치의 특허발명을 특허권자 허락 없이 실시하여 생성하거나 수집된 데이터를 구입하여 업으로 실시하는 행위, 즉 특허침해 행위로 생성되거나 수집된 데이터를 구입하고, 그 데이터의 속성 및 특성 등의 정보를 분석하는 행위(데이터마이닝), 그 데이터마이닝을 통하여 AI학습용 데이터구조를 가공하는 데이터전처리 행위에도 특허권의 효력이 미치는 것으로 해석해야 한다.

2. 데이터마이닝과 특허권의 효력제한

(1) 데이터의 활용과 데이터마이닝

데이터는 그 자체만으로도 상거래의 대상이 되고 있지만, 데이터를 이용하여 새로운 비즈니스를 하기 위해서는 데이터의 속성이나 특성을 분석하고, 그 분석한 결과를 바탕으로 데이터구조를 가공하고, 그 데이터구조를 근거로 기계학습(딥러닝학습)을 하여 최적 학습완료모델을 도출하여야 한다. 이러한 학습완료모델은 비즈니스모델의 AI기반 SW에 해당하는 것으로 이를 실제 상품화 또는 제품화에 실시함으로써 경제적 이익을 창출하게 된다.

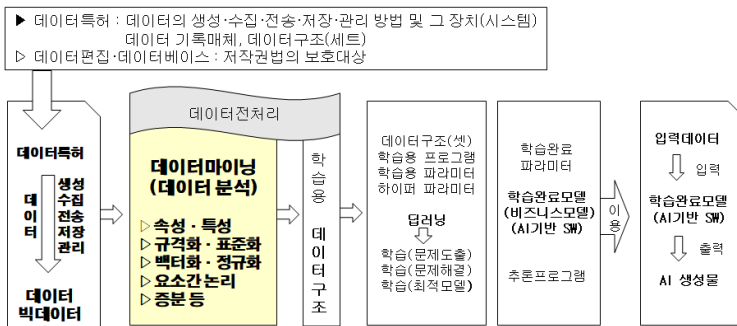
이처럼 최적의 비즈니스모델(AI기반 SW)을 개발하기 위해서는 반드시 데이터구조가 전제되어야 하고, 그 데이터구조를 가공하기 위해서는 데이터(원시데이터)가 있어야 한다. 데이터구조를 가공하기 위해서는 먼저 데이터의 분석이 이루어져야 한다. 즉, 데이터의 속성 또는 특성, 데이터의 규격화 또는

30) 특허청, 앞의 책(컴퓨터발명심사기준), 46-47면.

표준화, 데이터의 벡터화 또는 정규화, 데이터의 증분 등의 정보분석을 하여야 하며, 이러한 데이터의 정보분석을 데이터마이닝이라고 한다.

아래의<그림1>과 같이, 데이터마이닝은 AI딥러닝 학습에 필요한 데이터구조를 가공하는 행위가 아니라, 데이터구조를 가공하기 전(데이터전처리)에 데이터의 속성이나 특성 등을 분석하는 것을 말한다. 일반적으로 생성되거나 수집된 데이터의 속성이나 특성 등을 분석하고, 그 분석을 근거로 특정 목적 달성을 위한 AI학습용 데이터구조를 가공한다. 여기서 특정 과제 해결을 위한 AI학습용 데이터구조의 가공은 업으로서의 실사가 명확해지는 것이지만, 그러하지 않고 데이터 자체에 내재하여 있는 데이터의 규격 또는 표준화, 벡터화, 정규화, 증분 등의 속성이나 특성을 분석하는 데이터마이닝을 할 수 있다. 이러한 데이터마이닝의 결과를 바탕으로 특정 과제 해결을 위한 데이터구조가 가공되고, 그 데이터구조를 이용하여 AI학습에 의하여 비즈니스모델(AI기반 SW)이 개발되는 것이다. 비즈니스모델(AI기반 SW)을 “인간의 정신활동이 개입되지 않고 특정 과제 해결을 위하여 정보의 연산 및 가공을 실현하도록 하드웨어와 SW가 결합한 구체적 수단 또는 단계”³¹⁾로 청구범위에 기재한 경우, 그 출원발명은 발명의 성립성을 만족하여 특허를 받을 수 있다.³²⁾

<그림1 데이터의 활용과 데이터마이닝의 관계>



31) 특허청, 앞의 책(AI발명심사기준), 1206-1207면.

32) Shlomit Yanisky Ravid & Xiaoqiong (Jackie) Liu, “When artificial intelligence systems produce inventions: An alternative model for patent law at the 3A era”, *Cardozo L. Rev.*, Vol.39 (2018), pp.2221-2222.

(2) 데이터마이닝의 특허권 효력제한 여부

UR/TRIPs 협정에는 각 회원국에 특허권자의 적법한 이익 및 특허발명의 통상적 이용을 해하지 않는 조건으로 제3자의 적법한 이익을 고려하여 특허권의 효력을 제한할 수 있는 규정을 허용하고 있다(TRIPs 제30조). 이와 관련 현행 우리 특허법에 의하면, ‘연구 또는 시험’의 목적으로 특허발명을 실시하는 경우에는 특허권의 효력이 미치지 않는다고 규정하고 있다(제96조 제1항). 그리고 ‘연구 또는 시험’에는 “약사법에 따른 의약품의 품목허가·품목신고 및 농약관리법에 따른 농약의 등록을 위한 연구 또는 시험도 포함한다.”고 규정하고 있다(제96조 제1항 제1호).³³⁾ 다만, 특허권의 효력이 제한되는 ‘연구 또는 시험’의 범위에 대해서는 별도로 정의하고 있지 않고, 하위법령에도 이와 관련된 사항이 규정되어 있지 않다.

여기서 ‘연구 또는 시험’의 범위를 어디까지 볼 것인가가 중요하다. 특허권자의 허락 없이 특허발명을 실시한 행위에 특허권의 효력이 미치지 않는 것으로 보기 위해서는 객관적으로 특허발명 그 자체의 실시를 대상으로 해야 하며, 주관적으로는 특허요건의 실증이나 개량기술의 개발 등을 목적으로 해야 한다.³⁴⁾ 그렇지 않고 특정 목적이나 상업적 목적으로 특허발명의

33) 개정의 이유를 보면, 의약품의 품목허가·품목신고를 목적으로 의약특허발명의 시험이나 연구를 하는 경우 또는 시약품을 만드는 행위는 동일한 성분의 복제약(제네릭, 바이오 의약품)을 생산, 판매라는 상업적 목적을 가지고 ‘연구 또는 시험’을 하는 것이므로 “연구와 시험”에 해당하지 아니하였다. 그러나 다국적 제약사들은 하나의 의약품에 물질의약특허, 의약의 용도특허, 제형특허, 의약용도특허(투여용량·방법·시기를 구성요소로 보는 물질특허)를 근거로 값싼 제네릭 제품의 시장참여를 장기간 제한함으로써 국민건강 증진을 오히려 저해하는 문제점이 있었고, 그 문제점을 해결하기 위한 방안으로 의약품의 품목허가·품목신고를 목적으로 의약특허발명의 시험이나 연구(시약품 포함)는 제96조 제1항 제1호의 “연구 또는 시험”에 포함되는 것으로 개정하였다. 이에 대한 선행 논문으로는 김창화, “실험적 사용에 의한 특허권 제한 범위의 고찰- 미국법과의 비교를 중심으로”, 『財産法研究』, 제34권 제1호(2017), 160-165면; 박영규, “연구 또는 시험을 하기 위한 특허발명 실시의 의미와 한계”, 『산업재산권』, 제31호(2010), 25-34면; 조영선, “研究·試驗을 위한 特許發明의 實施와 特許權의 效力”, 『저스티스』, 통권 제116호(2010), 42-46면; 홍석경, “리서치용 특허의 이용에 관한 연구”, 『지식재산연구』, 제10권 제3호(2015), 40-43면 참조.

34) 윤선희, 『특허법』, 제6판, 법문사, 2019, 594면; 中山信弘·小泉直樹, 『新·注解 特許

‘시험 또는 연구’를 하는 경우에는 특허권 효력이 제한되지 않는 것으로 보는 것이 통설이다.³⁵⁾ 다만, 한미 FTA 협정 전에 약사법에 의한 의약품 품목 허가를 받을 목적으로 특허발명의 시약품 제조 및 그 시약품을 식품의약품 안전처에 신청하는 행위가 당해 특허권을 침해하는 것인지가 문제가 되었고, 이러한 침해 여부의 문제점을 해결하기 위한 방안으로 의약품의 품목허가 및 농약의 등록허가라고 하는 특정 목적에 필요한 ‘연구 또는 시험’에 대해서는 예외적으로 특허권의 효력이 제한되는 것으로 특허법을 개정하였다.

이러한 관점에서 보면, 데이터생성방법특허 및 그 장치특허, 데이터수집 방법특허 및 그 장치특허의 실시에 의하여 얻어진 데이터의 속성이나 특성 등을 분석하는 데이터마이닝이 특허법 제96조 제1항 제1호에서 규정하는 ‘연구 또는 시험’의 범위에 포함되어 특허권의 효력이 제한되는지가 명확하지 않다.

제4차 산업혁명 시대에서는 데이터라는 정보재가 없으면 AI학습을 할 수 없기 때문에 데이터에 내재한 속성이나 특성을 분석하는 데이터마이닝이 매우 중요하다. 이는 생성 및 수집된 데이터 그 자체만으로는 AI학습에 적용할 수 없기 때문에 데이터마이닝이 전제되어야 하기 때문이다. 만약 제3자가 특허권자의 허락 없이 특허발명의 실시에 의하여 생성되거나 수집된 데이터의 속성이나 특성을 분석하는 데이터마이닝을 한 경우, 그 데이터마이닝은 특허법상의 ‘연구 또는 시험’의 범위에 해당하는 것으로 볼 수 있는 것인지가 쟁점이 되고 있다.

3. 주요국 특허법상 특허권의 효력제한

(1) 미국

미국에 있어서 특허권의 효력이 제한되는 경우는 특허법 및 판례법으로 구분하여 검토할 수 있다.

法(中卷)』, 第2版, 青林書院, 2017, 1179-1180頁.

35) 송영식·이상정·황중환, 『지적소유권법(상)』, 육법사, 2008, 515면.

첫째, 미국 특허법에는 ‘시험 또는 연구’를 위한 타인의 특허발명을 실시하는 경우 특허권 효력이 제한된다는 명문규정이 없다. 다만, 미국 특허법 제271조(e)(1)³⁶⁾에는 미국식품의약국(FDA)에서 의약품의 제조허가를 받는데 필요한 정보 및 자료를 얻기 위해 타인의 특허발명을 실시(개발 및 정보 제출과 합리적으로 관련된 용도만으로 실시)하는 것은 특허침해가 아니라고 규정하고 있다. 본 규정은 정보 및 자료를 얻기 위한 특허발명의 실시에 한정하여 특허권의 효력이 미치지 않는 것이고, 만약 특허발명의 실시에 의하여 얻은 정보 및 자료를 이용하여 FDA에 의약품의 제조허가를 신청하는 행위는 동조 (e)(2)³⁷⁾에 의해 특허권 침해가 성립한다.³⁸⁾ 따라서 현행 미국 특허법상 데이터 생성·수집특허에 의해 얻어진 데이터를 특정 목적을 가지고 분석하는 데이터마이닝은 특허권의 효력이 제한되는 것으로 해석할 수 없다.

둘째, 미국 특허법에는 명문화되어 있지 않지만, 판례에 의하면 특허권의 효력이 제한되는 ‘시험 또는 연구’의 허용범위는 매우 제한적으로 인정하고 있다.³⁹⁾ FDA의 의약품 제조허가를 제외한 모든 기술분야의 특허권에 대한 특허발명의 ‘시험 또는 연구’를 위한 허용범위는 ‘시험 또는 연구’의 목적 및 범위에 의하여 결정되고 있다. 즉, 미국 판례상 ‘시험 또는 연구’로 인정되는 허용범위의 예로서는 통상적으로 특허발명의 정보를 얻을 목적으로 ‘시험

36) 미국 특허법 제271조(e)(1) (“특허발명(단, 새로운 동물 약품 또는 수의 생물제재(해당 용어는 「연방식품의약품화장품법」(Federal Food, Drug, and Cosmetic Act) 및 「1913년 3월 4일 법」(일반적으로 「Virus-Serum-Toxin Act」라 함)에서 사용되는 바와 같음)로, 재조합 DNA, 재조합 RNA, 하이브리도마 기술, 또는 위치 특이적 유전자조작기술 등의 기타 과정을 이용하여 주로 제조되는 것을 제외함)을 의약품 또는 수의 생물제재의 생산, 사용, 또는 판매를 규제하는 연방 법률에 따른 개발 및 정보 제출과 합리적으로 관련된 용도만으로, 미국에서 생산, 사용, 판매청약 또는 판매하거나, 또는 미국으로 수입하는 것은 침해행위로 되지 않는다”).

37) 미국 특허법 제271조(e)(1)(A) (“다음의 서류를 제출하는 것은 침해행위로 된다. (A) 특허에서 청구된 약품 또는 특허에서 청구된 사용에 대한, 「연방식품의약품화장품법」 제 505조(j)에 따른 또는 동법 제505조(b)(2)에 기재된 바의 신청서”).

38) 따라서 미국에서는 개발 및 정보 제출과 합리적으로 관련된 용도만으로 실시(임상시험 등)는 허용되고 있지만, 그 실시의 결과를 근거로 FDA의 의약품의 제조허가 신청은 특허권 존속기간이 종료된 후에 가능하다.

39) 김창화, 앞의 논문, 159면; 박영규, 앞의 논문, 15면.

또는 연구'를 하는 경우,⁴⁰⁾ 순수한 과학적·철학적 관심이나 흥미, 호기심 등을 확인하는 차원에서 '시험 또는 연구'를 하는 경우, 특허명세서에 기재된 사항이 기재요건 및 발명개시의무에 따라 정확한 것인지의 여부를 확인하기 위한 차원에서 '시험 또는 연구'를 하는 경우, 특허발명의 실시 가능성과 유용성을 확인하기 위한 차원에서 '시험 또는 연구'를 하는 경우, 일부 견해로 영리기관이라고 하더라도 그 결과물을 판매하거나 판매에 제공하지 않고 단순히 상업적·비상업적으로 '시험 또는 연구'를 하는 경우⁴¹⁾ 등을 들 수 있다. 한편 '시험 또는 연구'의 허용범위로 인정되지 않는 예로서는 '시험 또는 연구'를 통해 생성된 제품이나 방법을 상업적(판매, 사용 등)으로 이용하는 경우, 시장참여로 경제적 이익을 얻을 목적으로 '시험 또는 연구'를 하는 경우,⁴²⁾ 특허권자의 실시허락(라이선스)을 받지 않고 상업적으로 '시험 또는 연구'를 하는 경우⁴³⁾ 등을 들 수 있다.

이와 같이 미국 판례에서는 특허권자가 본인의 특허발명을 상업적 목적으로 이용하는 것과 같이, 특허발명의 '시험 또는 연구'를 하고 그 결과물을 이용하여 제품을 생산하여 판매하거나 그 방법을 사용하려는 제3자의 행위가 명백한 경우, 그 특허발명의 '시험 또는 연구'는 특허권 효력이 제한되지 않는다. 이러한 제3자의 '시험 또는 연구'는 이윤추구와 연계되어 있기 때문에 특허권자의 이익을 해할 염려가 있으므로 특허권 침해가 인정된다. 같은 논리로 데이터 생성·수집방법에 의하여 얻어진 데이터의 속성이나 특성을 분석하는 데이터마이닝이 미국 판례법상의 '시험 또는 연구'에 해당하여 특허권의 효력이 미치지 않는 것으로 해석할 수 있는 것인지는 명확하지 않다. 다만, 데이터마이닝의 목적이 특정 목적이나 과제 해결을 위한 AI학습용 데

40) Imperial Chemical Industries, PLC v. Henkel Corp., 545 F.Supp. 635 (D.C.Del. 1982).

41) David L. Parker, "Symposium, Patent Infringement Exemptions for Life Science Research", *HOUS. J. INT'L L.*, Vol.16 (1994), pp.659-660.

42) Richard E. Bee, "Experimental Use as an Act of Patent Infringement", *J. PAT OFF. SOC'Y*, Vol.34 (1957), p.375.

43) Jordan Karp, "Note, Experimental Use as Patent Infringement: The Impropriety of a Broad Exemption", *YALE L.J.*, Vol.100 (1991), p.2188.

이더구조를 가공하기 위한 ‘시험 또는 연구’로 보는 경우, 그 데이터마이닝은 상업적 목적의 이윤추구와 연계되어 있기 때문에 특허권자의 이익을 해할 염려가 있는 것으로 특허침해가 인정될 수 있다.

(2) 일본

일본 특허법에는 특허권의 효력이 미치지 아니하는 범위에 대하여 “특허권의 효력은 시험 또는 연구를 위한 특허발명의 실시에는 미치지 아니한다.”고 규정하고 있다(제69조 제1항). 다만, 일본 특허법에는 의약품의 품목허가·품목신고 및 농약의 등록허가를 위한 ‘연구 또는 시험’에는 특허권의 효력이 제한된다는 우리 특허법(제96조 제1항 제1호)과 같은 면책규정이 없다.

특허권의 효력이 제한되는 ‘시험 또는 연구’의 대상과 범위는 일본 특허법에 명문화하고 있지 않지만, 그 ‘시험 또는 연구’의 해당성 판단기준은 대상의 한정과 목적의 한정 요건이 있다. 대상의 한정 요건이란 특허발명 그 자체를 대상으로 해야 하고, 목적의 한정 요건이란 “기술의 진보”를 목적으로 하고 있다.⁴⁴⁾ 이와 같이 일본에서는 “기술의 진보”를 목적으로 하지 않고 특정 목적이나 상업적 목적으로 특허발명을 시험하거나 연구하는 경우에는 특허권 효력이 제한되지 않는다는 것이 통설이다.⁴⁵⁾

특히 일본에서는 데이터에 내재한 “데이터기술의 진보”를 목적으로 하는 데이터마이닝을 하는 경우 특허권의 효력이 제한되지만, 특정 과제 해결을 위한 AI학습용 데이터구조의 가공을 목적으로 데이터마이닝을 한 경우, 그 데이터마이닝은 “기술의 진보” 목적이 아니라 상업적 이용을 목적으로 한 것이므로 일본 특허법 제69조 제1항의 규정에 의한 특허권의 효력제한이 되지 아니하여 특허침해가 성립하는 것으로 해석할 수 있다. 따라서 데이터 생성·수집특허에 의하여 얻어진 데이터의 속성이나 특징을 분석하는 데이터마이닝이 일본 특허법 제69조 제1항의 ‘시험 또는 연구’에 해당하는 것으로 볼 수 있는 것인지에 대한 명확한 해석 증거는 없다. 다만, 일본 정부는 특허권

44) 中山信弘·小泉直樹, 앞의 책, 1179頁.

45) 中山信弘·小泉直樹, 위의 책, 1179-1186頁.

에 의하여 얻어진 데이터의 여부와 관계없이 ‘생성·수집데이터’를 하나의 정보재로 간주하고, 그 ‘생성·수집데이터’를 정보재로서의 이용과 활용에 있어서는 데이터의 경제적 이용에 관한 당사자 계약에 의하여 이루어질 수 있도록 데이터이용계약가이드라인을 기업에 제공하고 있다.⁴⁶⁾

다만, 일본 저작권법에는 ‘생성·수집된 데이터’에 타인의 저작물이 있는 경우, 그 ‘생성·수집데이터’의 속성이나 특성을 분석하는 데이터마이닝에는 저작권의 효력을 제한하는 규정이 도입되었다.⁴⁷⁾ 즉, 저작물에 표현된 사상 또는 감정을 자신이 향수(享受)하지 않거나 타인에게 향수하도록 하지 않는 ‘비향수(非享受)의 이용’에 있어서, ‘생성·수집된 데이터’의 정보 해석(정보의 추출, 비교, 분류, 기타 해석, 분석을 실시하는 것) 및 그 정보 해석의 용도로 제공하는 행위에는 저작권의 효력이 제한된다는 면책규정⁴⁸⁾을 도입하였다.⁴⁹⁾⁵⁰⁾

46) 經濟産業省 情報經濟課, 앞의 책, 21-28頁.

47) 2018. 5. 25. 공포, 2019. 1. 1. 시행 일본 저작권법 34조의4.

48) 일본 저작권법 제34조의4(“저작물은 다음의 경우 또는 기타 해당 저작물에 표현된 사상 또는 감정을 자신이 향수(享受)하거나 타인에게 향수하도록 하는 것을 목적으로 하지 않는 경우에는 필요하다고 인정되는 한에서 어떠한 방법으로도지 이용할 수 있다. 다만 해당 저작물의 종류 및 용도와 해당 이용 형태에 비추어 저작권자의 이익을 부당하게 침해하는 경우에는 그러하지 아니하다. 1) 저작물의 녹음, 녹화, 기타 이용에 관한 기술의 개발 또는 실용화를 위한 시험의 용도로 제공하는 경우, 2) 정보 해석(다수의 저작물 또는 기타 대량의 정보로부터 해당 정보를 구성하는 언어, 음, 영상 기타 요소에 관한 정보를 추출, 비교, 분류, 기타 해석(분석)을 실시하는 것을 말한다. 제47조의5 제1항 제2호와 같다)의 용도로 제공하는 경우, 3) 앞의 제1호와 제2호의 경우 이외에 저작물의 표현에 대해 인간의 지각에 의한 인식을 수반하지 않고 해당 저작물을 컴퓨터에 의한 정보처리 과정에서 이용 또는 그 밖의 이용(프로그램저작물에 있어서는 해당 저작물의 컴퓨터에서의 실행을 제외한다)에 제공하는 경우”).

49) 데이터마이닝에 관한 일본 저작권법의 개정 내용은 김성원·최상민·이수원, “텍스트·데이터 마이닝 과정의 저작물 이용 면책 규정 신설안에 대한 소고-적대적 생성신경망에의 적용”, 『지식재산연구』, 제18권 제1호(2023), 153-156면; 차상욱, “저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점—텍스트·데이터마이닝(TDM) 면책규정”, 『경영법률』, 제32집 제1호(2021), 25-27면 참조.

50) 우리 저작권법에서 일본의 면책규정과 같은 취지로 저작권 제35조의5(2020. 11. 2.) 또는 제43조(2021. 1. 15.)의 개정안이 각각 심의되었지만, 국회를 통과하지 못하였다. 이와 관련한 논문으로는 차상욱, 위의 논문, 29-43면; 박성호, “저작권법 전면 개정안 주요 내용의 비판적 고찰”, 『정보법학』, 제24권 제3호(2020), 12-17면; 이상용, “데이터의

(3) 독일

독일 특허법 제11조 제2항에 의하면, 유럽공동체특허협약(Community Patent Convention) 제31(b)에 근거하여 “특허발명의 내용에 관해 시험 목적으로 행하는 행위”에 대해서는 특허권의 효력이 미치지 않는다고 규정하고 있다.⁵¹⁾ 여기서 ‘특허발명의 내용’에 관한 시험 목적이므로, 특허발명의 특성이나 기능, 작동원리 등에 대한 시험이나 연구가 대상이 된다. 따라서 특허발명에 내재한 원리나 사실을 확인하거나 검증하는 시험 또는 특허발명보다 개량되거나 진보된 기술 사상을 도출하기 위한 시험, 무효심판 등의 특허분쟁에 활용하고자 특허요건을 분석하기 위한 특허발명의 시험 등에는 특허권 효력이 제한되는 것이다.⁵²⁾ 물론 이러한 특허발명의 시험이 기술의 진보발전을 촉진한다거나 국민경제의 발전 또는 공중의 이익, 국민건강 증진 등을 목적으로 하는 경우에도 특허권 효력이 제한되는 것이다.⁵³⁾ 따라서 독일에서는 특허발명의 시험으로 얻어진 기술지식을 특정 목적으로 이용하는지 여부와 관련 없이, 당해 특허발명에 내재한 특성이나 기능, 작용 효과 등과 같은 기술지식을 얻을 목적으로 하는 시험이 포함되는 것으로 해석할 수

비계약적 이용-데이터 마이닝을 위한 저작권 제한을 중심으로, 『강원법학』, 제65권(2021), 14-17면; 홍승기, “데이터마이닝 면책 입법 방향에 대한 의문”, 『경영법률』, 제32집 제4호(2022), 4-11면.

- 51) 독일 특허법 제11조에는 특허의 효력이 미치지 않는 범위를 다음과 같이 규정하고 있다.
1. 비상업적 목적으로 사적으로 행하는 행위
 2. 특허발명의 내용에 관해 실험 목적으로 행하는 행위
 - 2a. 식물신품종의 육성, 발견 및 개발을 위한 생물학적 재료의 사용
 - 2b. 의약품을 유럽연합 시장에 투입하는 판매허가 및 유럽연합가맹국 또는 제3국에서 의약품 판매승인을 취득하기 위해 필요한 연구, 시험 및 그 후의 실무적 요건
 3. 의사의 처방에 의해 약국 내에서 개별적으로 이루어지는 의약의 조합 또는 이와 같이 조합된 의약에 관한 행위
 4. 산업재산권 보호에 관한 파리조약의 다른 회원국의 선박이 일시적 또는 우발적으로 동법의 효력이 적용되는 영해에 들어온 경우
 5. 산업재산권 보호에 관한 파리조약의 다른 회원국의 항공기 또는 차량이 일시적 또는 우발적으로 동법의 시행 영역에 들어온 경우, 항공기 또는 차량의 구조 또는 조작 또는 해당 항공기 또는 차량의 부속물에 있어서 특허발명의 사용
- 52) 박영규, 앞의 논문, 36-37면.
- 53) Metzger, Patentverletzung durch Forschung, GRUR Int. 1967, S. 124, 126.

있다.

같은 취지로 데이터 생성·수집방법특허발명의 실시에 의하여 얻어진 데이터의 속성이나 특성을 분석하는 데이터마이닝은 독일 특허법상의 ‘특허발명의 내용에 관한 시험 목적’의 범위에 포함되는 것인지는 명확하지 않다. 다만, 특허발명 그 자체의 실시가 아니라 그 특허발명의 실시에 의하여 생성된 데이터의 속성이나 특성을 분석하는 데이터마이닝에는 특허권의 효력이 미치지 않지만, 상업적으로 이용하기 위해 당해 데이터의 속성이나 특성을 분석하는 데이터마이닝은 독일 특허법 제11조 제2항의 규정이 적용되지 아니하여 특허권의 효력이 제한되지 않는 것으로 해석할 수 있다.

(4) 중국

중국 특허법 제75조 제4항에 의하면 “과학연구 및 시험만을 위하여 관련 특허를 실시하는 행위”에는 특허권의 효력이 제한된다고 규정하고 있고, 동조 제5항에는 “행정적 심사·허가에 필요한 정보를 제공하기 위하여 특허약품 또는 특허의료기계를 제조, 사용, 수입하는 경우 및 특별히 이를 위하여 특허약품 또는 특허의료기계를 제조, 수입하는 행위”에도 특허권의 효력이 제한되는 것으로 규정하고 있다.⁵⁴⁾ 이와 같이 중국은 특허권의 효력이 제한되는 범위를 ‘과학연구 및 시험’만으로 한정하고 있으므로, 실질적으로 특허발명의 내용에 대하여 과학적 시험을 하는 경우에만 특허권의 효력이 제한되는 것으로 엄격하게 해석하고 있다. 특허발명에 내재한 원리나 사실을 확

54) 중국 특허법 제75조에서 규정하는 “특허권 침해로 보지 아니하는 행위”는 다음과 같다.

1. (생략)
2. (생략)
3. 중국의 영토, 영해, 영공을 임의로 통과하는 외국의 운송수단이 그 소속 국가와 중국이 체결한 협약 또는 공동으로 가입한 국제조약에 따라서 또는 호혜원칙에 따라서 운송수단 자체의 수요를 위하여 그 장치 및 설비 중에 관련 특허를 사용하는 행위
4. 과학연구 및 시험만을 위하여 관련 특허를 사용하는 행위
5. 행정적 심사·허가에 필요한 정보를 제공하기 위하여 특허약품 또는 특허의료기계를 제조, 사용, 수입하는 경우 및 특별히 이를 위하여 특허약품 또는 특허의료기계를 제조, 수입하는 행위

인하거나 검증하는 시험 또는 연구에는 특허권의 효력이 제한되지만, ‘과학 연구 및 시험만’의 범위에 포함되지 않고 특허발명으로부터 개량발명이나 진보된 기술 사상만을 연구하는 행위, 특정 목적하에서 시험을 하는 경우에는 특허권의 효력이 제한되지 않는 것으로 해석할 수 있다.

같은 취지로 데이터 생성·수집방법특허발명의 실시에 의하여 얻어진 데이터의 속성이나 특성을 분석하는 데이터마이닝은 중국 특허법상의 ‘과학연구 및 시험만을 위하여 관련 특허를 실시하는 행위’의 범위에 포함되는 것인지는 명확하지 않다. 데이터특허권 그 자체의 실시가 아니라 그 특허발명의 실시에 의하여 생성된 데이터의 속성이나 특성을 분석하는 데이터마이닝은 특허권의 효력이 제한된다. 그러나 특정 과제 해결을 위한 AI학습용 데이터구조를 가공하기 위한 목적을 가지고 데이터마이닝을 하였다면, 그 데이터마이닝은 상업적인 이용을 목적으로 한 것이므로 중국 특허법 제75조 제4항의 규정이 적용되지 아니하여 특허권의 효력이 미치게 되어 특허침해가 성립하는 것으로 해석할 수 있다.

(5) 소결

미국 및 일본, 독일, 중국에는 표현상 약간의 차이가 있지만, 우리 특허법과 같이 특허권의 효력이 제한되는 ‘시험 또는 연구’와 같은 취지의 규정을 두고 있다. 다만, ‘시험 또는 연구’의 범위와 그 판단기준을 특허법에 명문으로 규정하고 있지 않고, 판례 또는 각국 심사기준으로 보충 설명하고 있다. 또한 ‘연구 또는 시험’에 의약품의 품목허가를 위한 연구나 시험 및 그 품목허가를 위한 시약품에 대해서는 특허권의 효력이 미치지 않는다는 규정을 둔 국가가 있는 반면, 일본과 같이 명문화하지 않은 경우도 있다. 그러나 이들 모든 국가의 특허법에는 데이터마이닝은 “시험 또는 연구”에 포함되어 특허권의 효력이 제한된다는 명문의 규정은 없다. 다만, 데이터마이닝은 통상의 ‘시험 또는 연구’에 해당하는 것인지를 판단하고 있고, 특히 데이터특허의 기술적 특성, 기술적 개량, 진보 등을 파악하기 위한 시험·연구의 경우에 한정해서 특허권의 효력을 제한하고 있다. 특정 목적이나 과제 해결을 위한

AI학습용 데이터구조를 가공하기 위한 데이터마이닝은 상업적 목적을 가지고 있다는 이유로 특허권의 효력이 제한되지 않는 것으로 보고 있다. 다만, 데이터마이닝이 특정 목적이나 과제 해결을 위한 AI학습용 데이터구조를 가공하기 위한 것인지의 여부를 판단하는 기준이 명확히 제시되어 있지 않는 것으로 판단된다.

IV. 데이터마이닝을 통한 데이터의 이용 극대화 방안

1. 시험·연구에 데이터마이닝을 포함하는 방안

데이터의 이용 극대화를 위한 방안으로 먼저 제3자가 자유롭게 데이터마이닝을 할 수 있는 방안을 검토하여 본다. “연구 또는 시험”을 위한 특허발명의 실시는 특허권의 효력이 미치지 않는다는 현행 특허법 규정(제96조 제1항 제1호)에 있어서, “연구 또는 시험”이란 특허발명 자체를 대상으로 연구나 시험을 하는 것을 의미하고, 그 특허발명을 다른 사용 목적이나 상업적 사용을 위한 하나의 수단으로 하는 연구나 시험은 제외된다고 해석하는 것이 통설이다.⁵⁵⁾ 이러한 관점에서 본다면, ‘데이터 생성·수집방법특허’의 특성·속성·신뢰성, 기술성·실현 가능성 및 그 기술성·실현 가능성의 향상이나 개량 등을 위한 시험·연구는 “연구 또는 시험”의 범위에 포함되므로 특허권의 효력이 제한된다.

여기서 쟁점이 되는 것은 특허발명 그 자체의 시험이나 연구를 하는 것이 아니라, 방법·장치특허발명의 실시에 의하여 얻어진 생성물에 대하여 시험이나 연구를 하는 경우에도 특허권의 효력이 제한되는지가 명확하지 않다는 것이다. 예를 들면, ‘데이터 생성·수집방법특허발명’의 실시에 의하여 얻어진 데이터의 속성이나 특성을 분석하는 데이터마이닝이 “연구 또는

55) 송영식·이상정·황중환, 앞의 책, 515면.

시험”의 범위에 포함되어 특허권의 효력이 제한되는 것인지, 아니면 데이터를 방법특허에 의하여 생성된 물건 또는 매체에 저장된 물건으로 보고 생산방법의 추정 규정(제129조)에 의거하여 특허침해가 성립하는 것인지가 쟁점이 된다.⁵⁶⁾

물론 특허발명의 실시에 의하여 생성하고 수집한 데이터는 정보재로서 상거래의 대상이 되지만, 그 데이터에 내재한 데이터의 속성이나 특성이 무엇인지를 분석하는 데이터마이닝이 “연구 또는 시험”에 속하는 것으로 해석하여 특허권의 효력이 제한되는 것으로 볼 수 있는지가 명확하지 않다. ‘데이터 생성·수집방법특허’의 적법한 실시에 의하여 창출된 데이터가 있고, 그 데이터의 속성이나 특성을 분석하는 제3자의 행위, 즉 제3자의 데이터마이닝은 특허발명의 실시에 해당하여 특허침해가 성립하는 것인지가 쟁점이 될 수 있다.

제4차 산업혁명 시대에 있어서 신산업(AI기반 기술) 또는 새로운 비즈니스 모델(AI기반 SW)을 창출하기 위해서는 데이터가 반드시 요구되고 있다. 이러한 데이터의 생성과 수집에 관한 기술개발자의 특허보호와 그 데이터의 활용을 목적으로 하는 데이터분석자의 연구 활동 보호가 형평성하에서 상호간 조화를 꾀할 수 있는 제도의 도입이 필요하다. 특허권자는 데이터의 생성과 수집방법 연구에 투자한 연구비의 회수와 이익 창출을 도모할 수 있도록 하는 특허권의 효력 범위가 필요하고,⁵⁷⁾ 한편으로는 당해 특허권에 의한 산업발전을 촉진한다는 특허법의 목적에 따라 데이터의 속성이나 특성을 분석하는 데이터마이닝을 누구나 자유롭게 할 수 있도록 하는 특허권의 효력제한 규정이 필요하다.

데이터특허권자는 특허발명의 실시에 의하여 얻어진 데이터의 사용을 통하여 연구개발비 회수 및 경제적 이익을 창출할 수 있다. 이를 위해서는 당

56) 이와 관련한 판례를 검색하였지만, 찾을 수 없었다.

57) 데이터 생성·수집특허권자는 그 특허발명의 실시에 의하여 수집된 데이터를 근거로 AI기반 비즈니스모델 사업을 실시하거나, 그 데이터를 다른 사람에게 판매하여 연구개발비를 회수할 수 있다. 이러한 관점에서 보면, 제3자의 데이터마이닝은 결국 AI학습용 데이터구조를 가공하기 위한 목적을 전제로 할 수 있으므로 특허침해가 성립할 수 있다.

해 데이터의 사용이 적극적으로 이루어져야 한다. 그 사용을 위해서는 제일 먼저 데이터의 속성이나 특성을 분석하는 데이터마이닝이 있어야 그다음에 AI학습용 데이터구조가 가공되기 때문이다. 특허데이터구조를 가공하기 위해서는 데이터전처리 과정으로 데이터마이닝이 전제되어야 하므로, 그 데이터마이닝을 누구나 자유롭게 할 수 있어야 한다. 데이터구조의 가공행위는 특정 목적의 AI학습용으로 데이터 요소의 논리적 구조를 가공하는 것이므로 당연히 특허침해가 되지만, 그 전 단계인 데이터마이닝은 데이터전처리 과정의 일환으로 특정 목적을 가지고 이루어지는 것이 아니므로 누구나 자유롭게 할 수 있도록 하는 특허권의 효력제한 규정을 특허법에 도입하는 것을 제언하고자 한다.

의약품의 품목허가 및 농약 등록을 위한 특허발명의 실시(예, 시약품 제조 등)는 “연구 또는 시험”에 포함되기 때문에 특허권의 효력이 제한된다. 이러한 규정의 도입 취지는 특허권의 종료 전에 의약특허발명(물질, 용도, 제제, 의약품)의 실시를 하고, 그 특허권 종료 후 값싼 제네릭의약품이나 바이오의 약품을 시장에 조기에 출현하여 국민건강 증진을 촉진하기 위함이다.⁵⁸⁾

이와 같은 취지로 데이터의 속성이나 특성을 분석하고, 그 분석 결과를 활용하여 AI의 응용 분야를 확대해 산업발전을 촉진할 수 있도록 데이터마이닝을 “연구 또는 시험”에 포함되는 것으로 특허법을 개정할 필요가 있다. 구체적인 개정 방안으로는 아래의 <표1>과 같이 특허법 제96조 제1항 제1호에 ‘데이터의 속성·특성을 확인하기 위한 정보분석’이 “연구 또는 시험”의 범위에 포함되는 것으로 개정할 필요가 있다. 데이터마이닝에 특허권의 효력이 미치지 않는 것으로 특허법이 개정되면, 누구나 데이터특허발명의 속성이나 특성 등의 정보분석을 자유롭게 실시할 수 있음으로써 보다 진보되고 개량된 데이터발명의 연구개발(기술의 진보)을 촉진할 수 있을 것이다.

58) 김창화, 앞의 논문, 159면; 박영규, 앞의 논문, 15면.

〈표1 데이터마이닝을 특허권 효력제한 규정의 도입 방안〉

현행(특허법)	개정안
제96조(특허권의 효력이 미치지 아니하는 범위) ① 특허권의 효력은 다음 각 호의 어느 하나에 해당하는 사항에는 미치지 아니한다. 1. 연구 또는 시험(「약사법」에 따른 의약품의 품목허가·품목신고 및 「농약관리법」에 따른 농약의 등록을 위한 연구 또는 시험을 포함한다)을 하기 위한 특허발명의 실시 2. (생략) 3. (생략) ② (생략)	제96조(특허권의 효력이 미치지 아니하는 범위) ① 특허권의 효력은 다음 각 호의 어느 하나에 해당하는 사항에는 미치지 아니한다. 1. 연구 또는 시험(「약사법」에 따른 의약품의 품목허가·품목신고 및 「농약관리법」에 따른 농약의 등록을 위한 연구 또는 시험, 데이터의 속성·특성을 확인하기 위한 정보 분석을 포함한다)을 하기 위한 특허발명의 실시 2. (생략) 3. (생략) ② (생략)

2. 생산방법 추정의 물건에 데이터를 포함하는 방안

데이터의 이용 극대화를 위하여 제3자의 데이터마이닝이 특허권 효력이 미치지 않도록 하는 효력제한 규정을 도입하는 경우 특허권자의 권리행사가 제한되므로, 이에 대응하여 데이터수집방법에 의하여 수집된 데이터를 물건으로 간주하고 그 물건의 생산 및 양도 등에도 특허권의 효력이 미치도록 특허법의 개정 방안을 검토하여 본다.

방법특허의 실시범위는 그 방법의 사용 또는 그 방법 사용의 청약⁵⁹⁾을 말한다. 방법특허가 물건을 생산하는 방법발명인 경우에는 “그 물건과 동일한

59) 데이터방법특허에서 “그 방법의 사용을 청약하는 행위”도 실시의 범위에 포함되고, 이 경우 특허권의 효력은 그 방법의 사용이 특허권(전용실시권 포함)을 침해한다는 것을 알면서 그 방법의 사용을 청약하는 행위에도 미친다(특허법 제94조 제2항). 따라서 데이터 생성방법이 프로그램발명인 경우, 제3자가 특허발명인 데이터생성프로그램을 인터넷을 통하여 전송하는 행위는 실제 사용을 하지 않았지만, 그 사용의 청약을 하였으므로 당해 특허권의 인지 요건을 전제로 침해가 성립하게 된다.

물건”이 특허출원 전에 국내에서 공지·공연실시된 물건, 국내·국외에서 반포된 간행물에 게재된 물건 및 전기통신회선상 이용할 수 있는 물건이 아니면 그 방법특허에 의하여 생산된 것으로 추정하고 있다.⁶⁰⁾ 만약, ‘데이터 생성·수집방법특허’에 의하여 생산된 데이터를 물건으로 보는 경우 또는 매체에 저장된 물건으로 보는 경우에는 그 데이터에도 특허권의 효력이 미치게 된다. 즉 데이터를 물건으로 보는 경우, 다른 사람이 그 데이터를 ‘사용, 양도, 대여, 수입, 양도 또는 대여를 위한 청약’을 하는 행위를 하게 되면 당해 방법특허권을 침해하게 된다.

‘데이터 생성·수집방법특허’에 의하여 생산된 데이터는 정보재로서 매체에 저장된 형태로 보관·관리되고 있고, 데이터의 상거래에서도 매체에 저장된 정보재로 이루어지고 있으므로 물건의 유형으로 볼 수 있을 것으로 판단된다. 다만, 현행 특허법에는 데이터를 물건으로 볼 수 있다는 명문의 규정이 없다. 따라서 데이터를 물건으로 간주하여 생산방법특허의 추정 규정을 적용할 수 없는 것으로 해석함이 타당하다. 대용량 데이터의 생성과 수집, 그 수집된 대용량 데이터의 전송과 저장·관리 등에는 고도의 기술이 수반되어야 한다. 예를 들면, 자율주행차의 AI기반 안전운행시스템을 개발하기 위해서는 특정 지역별 다양한 교통정보를 수시·비정형 형태로 생성·수집하는 기술, 그 수집된 대용량의 데이터를 관리센터로 전송하는 기술, 대용량 데이터를 저장·관리하는 기술 등이 수반되어야 하고, 이러한 데이터 관련 기술개발에는 많은 자금과 인력이 투자되어야 하므로 당해 데이터의 적극적 사용을 통하여 이익을 창출하여야 한다. 이와 같이 데이터산업뿐만 아니라 AI산업발전을 더욱 촉진하기 위해서는 데이터의 생성·수집이 적극적으로 이루어져야 하고, 그 데이터 생성·수집 등의 기술을 개발한 특허권자의 보호도 데이터의 이용 촉진과 함께 적절히 이루어질 수 있는 방안의 도입이 필요하다.

이와 같은 취지로 데이터를 물건으로 본다는 규정을 특허법에 반영하여 ‘데이터 생성·수집방법특허’의 효력범위를 확대하는 것을 제안하고자 한다. 방법특허에 의하여 생성된 데이터를 분석하는 데이터마이닝을 “시험 또는

60) 특허법 제129조 제1항 및 제2항.

연구”에 포함하여 특허권의 효력을 제한하는 것에 대응한 특허권자의 보호 강화 차원에서 당해 데이터를 생산방법의 추정 규정을 적용할 수 있도록 하는 제도를 도입하여야 한다. 즉, 물건을 생산하는 방법특허에 있어서 ‘생산방법의 추정’에 데이터를 물건으로 본다는 취지의 규정을 도입하고 그 데이터에도 특허권의 효력이 미치게 할 필요가 있다. 그 구체적인 개정 방안으로는 아래의 <표2>와 같이 특허법 제129조 본문에 ‘물건을 생산하는 방법의 발명’에서의 물건에 데이터가 포함된다는 것으로 개정할 필요가 있다.

<표2 생산방법 추정의 물건에 데이터를 포함하는 방안>

현행(특허법)	개정안
제129조(생산방법의 추정) 물건을 생산하는 방법의 발명에 관하여 특허가 된 경우에 그 물건과 동일한 물건은 그 특허된 방법에 의하여 생산된 것으로 추정한다. 다만, 그 물건이 다음 각 호의 어느 하나에 해당하는 경우에는 그러하지 아니하다. 1. (생략) 2. (생략)	제129조(생산방법의 추정) 물건(데이터를 포함한다. 이하 같다)을 생산하는 방법의 발명에 관하여 특허가 된 경우에 그 물건과 동일한 물건은 그 특허된 방법에 의하여 생산된 것으로 추정한다. 다만, 그 물건이 다음 각 호의 어느 하나에 해당하는 경우에는 그러하지 아니하다. 1. (생략) 2. (생략)

3. 데이터의 물건 인정 방안

데이터가 민법상 물건인가 아닌가는 별론으로 하고, 특허법에는 이와 관련 명문의 규정이 없고, 특허청의 컴퓨터발명심사기준에는 청구범위의 청구항 말미에 ‘데이터 기록매체’⁶¹⁾ 청구항을 기재할 수 있고,⁶²⁾ 이러한 청구항

61) 데이터 기록매체는 “기록된 데이터구조로 말미암아 컴퓨터가 하는 처리 내용이 특정되는 「구조를 가진 데이터를 기록한 컴퓨터로 읽을 수 있는 매체」를 말한다[특허청, 앞의 책(컴퓨터발명심사기준), 43-44면].

62) 데이터 기록매체 청구항은 다음과 같이 기재한 경우를 말한다.

“제1항 A 구조, B 구조, C 구조, …를 가진 데이터를 기록한 컴퓨터로 읽을 수 있는 매체”

은 물건발명으로 간주하여 특허를 받을 수 있다.⁶³⁾⁶⁴⁾ 반면 특허청의 AI발명 심사기준에는 ‘데이터구조 기록매체 청구항’을 물건발명으로 인정하는 판단 기준을 제시하고 있지만, ‘데이터 기록매체 청구항’을 물건발명으로 인정한다는 명문의 규정은 없다. 그러나 AI발명심사기준에서도 ‘데이터구조 기록매체 청구항’뿐만 아니라 ‘데이터 기록매체 청구항’을 물건발명으로 인정하는 것으로 해석할 수 있다.

청구범위에는 데이터 그 자체를 청구항 말미에 카테고리 기재할 수 없고, 반드시 데이터가 매체(컴퓨터 등)에 기록되거나 저장된 매체 형태의 카테고리로 기재하는 경우에 물건발명으로 볼 수 있다. 이러한 물건발명의 대상은 데이터 그 자체가 아니라 데이터가 기록되거나 저장된 매체를 말하는 것으로 데이터만을 물건발명으로 보호할 수 없는 문제점이 있다. 데이터와 유사한 객체로서 ‘프로그램’을 물건발명으로 인정한다는 규정이 특허법에는 명문화되어 있지 않지만, 특허청의 컴퓨터발명심사기준에는 ‘프로그램 기록매체 청구항’⁶⁵⁾뿐만 아니라 프로그램을 물건으로 인정하는 ‘프로그램 청구항’에 대해서도 특허를 받을 수 있다고 규정하고 있다. 즉, 청구범위의 청구항 말미에 ‘하드웨어와 결합한 「기록매체에 저장된 컴퓨터프로그램」’으로 기재하는 경우에는 물건발명으로 특허를 받을 수 있다. 하드웨어(기록매체)에 저장된 프로그램을 물건으로 보고 특허를 허여하고 있다. 이와 같이 우리나라에서는 특허법이 아니라 특허청 컴퓨터발명심사기준에 하드웨어(기록매체)에 저장된 프로그램을 물건발명으로 인정하고 있지만, 일본 특허법에서는

63) 특허청, 앞의 책(컴퓨터발명심사기준), 46-47면.

64) 예를 들면, 등록특허번호 10-0716242(등록일자: 2007. 5. 2., 특허권자: Sony), 발명의 명칭은 ‘데이터 기록매체, 기록장치 및 재생방법’이고 그 청구범위 제1항에는 “메인데이터와 재생경과시간을 … 있어서, … 기록될 상기 메인 데이터의 일부로서 수신되는 사용자데이터를 기록하기 위한 데이터 영역(data area)을 포함하여 구성된 것을 특징으로 하는 데이터 기록매체”라고 기재하고 있다.

65) 예를 들면, 등록특허번호 10-2519467(등록일자: 2023. 4. 4., 특허권자: 캠브리콘 테크놀로지스 코퍼레이션 리미티드)의 청구범위 제10항에는 “컴퓨터프로그램이 내장된 컴퓨터 판독 가능 저장매체로서, 상기 컴퓨터프로그램은 프로세서에 의해 실행될 경우, 제1항 내지 제8항 중 어느 한 항에 따른 데이터를 전처리하기 위한 방법의 단계를 수행하는 것을 특징으로 하는, 컴퓨터 판독 가능 저장매체”라고 기재하고 있다.

‘프로그램’을 물건에 포함한다고 규정되어 있지만,⁶⁶⁾ 데이터가 물건에 포함된다는 명문의 규정은 없다.

제4차 산업혁명 시대뿐만 아니라 미래의 산업혁명에서도 데이터가 정보재로서 매우 중요한 역할을 하고 있으므로, 데이터의 생성·수집, 전송·저장·관리에 관한 기술개발의 촉진이 필요하고, 그 기술개발의 성과물을 데이터 특허로 보호하기 위해서는 데이터를 물건으로 본다는 명문의 규정이 필요하다. 데이터를 물건으로 보는 경우에는 청구항의 말미에 「...데이터」로 기재할 수 있고, 그 데이터의 실시범위가 넓어짐으로써 특허권자의 적극적 보호가 이루어질 수 있다. 이는 특허권자의 보호만을 강구하는 것이 아니라, 위에서 제안한 데이터마이닝을 특허권의 효력이 미치지 않는 “시험 또는 연구”에 포함하는 것으로 특허법을 개정하는 경우에는 데이터 관련 발명의 특허권자 보호와 그 데이터이용자의 보호가 형평성 관점에서 상호 달성될 수 있기 때문이다.

개정 방안으로는 아래의〈표3〉과 같이 특허발명의 실시를 규정한 제2조 제3호의 “물건의 발명인 경우”에서, 물건의 범위에 데이터가 포함되는 것으로 개정하는 것을 제안하고자 한다. 즉 “물건의 발명인 경우”를 “물건(데이터를 포함한다)의 발명인 경우”로 개정함으로써 데이터특허발명의 보호를 강화하고, 한편으로는 데이터특허에 관한 데이터마이닝에는 특허권의 효력이 제한되는 것으로 개정하여 제3자의 자유로운 시험이나 연구를 할 수 있도록 하여 데이터산업 및 그 데이터이용에 의한 AI산업의 발전을 촉진할 수 있을 것으로 판단된다. 아래의 〈표3〉과 같이 개정되는 경우, 물건을 생산하는 방법 특허에서의 물건에도 데이터가 당연히 포함되기 때문에 위에서 검토한 생산방법의 추정 규정(제129조)은 개정할 필요는 없다.

66) 일본 특허법 제2조 제3호 1목에는 “물(物)(프로그램 등을 포함한다. 이하 동일)의 발명에는(物(プログラム等を含む。以下同じ))”이라고 규정하고 있다.

〈표3 데이터를 물건에 포함하는 방안〉

현행(특허법)	개정안
제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 1. (생략) 2. (생략) 3. “실시”란 다음 각 목의 구분에 따른 행위를 말한다. 가. 물건의 발명인 경우: 그 물건을 생산·사용·양도·대여 또는 수입하거나 그 물건의 양도 또는 대여의 청약(양도 또는 대여를 위한 전시를 포함한다. 이하 같다)을 하는 행위 나. (생략) 다. (생략)	제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 1. (생략) 2. (생략) 3. “실시”란 다음 각 목의 구분에 따른 행위를 말한다. 가. 물건(데이터를 포함한다. 이하 같다)의 발명인 경우: 그 물건을 생산·사용·양도·대여 또는 수입하거나 그 물건의 양도 또는 대여의 청약(양도 또는 대여를 위한 전시를 포함한다. 이하 같다)을 하는 행위 나. (생략) 다. (생략)

만약, 데이터가 물건인지의 여부에 대한 논쟁으로 <표3>과 같이 제2조 제3호를 개정할 수 없는 경우에는 ‘프로그램’을 물건으로 인정한 특허청의 컴퓨터발명심사기준과 같은 취지로 동 심사기준 및 AI발명심사기준에 “하드웨어와 결합한 「기록매체에 저장된 데이터」”를 물건발명으로 인정한다는 취지의 가이드라인을 도입할 필요가 있다.

V. 결 론

제4차 산업혁명 시대에서는 데이터가 반도체와 같이 산업의 쌀이고 석유에 해당하는 것으로 비유되고 있다.⁶⁷⁾ 최근 산업계에서는 데이터를 이용한 새로운 비즈니스모델에 의한 신사업의 촉진과 신제품의 개발을 적극 추진하

67) 박진아, “데이터의 보호 및 유통 법제 정립 방안”, 『서강법률논총』, 제9권 제2호(2020), 3면.

고 있다. 이러한 새로운 비즈니스모델은 데이터 또는 빅데이터를 기반으로 이루어지고 있기 때문에 데이터의 생성과 수집에 관한 기술개발이 활발하게 이루어지고, 그 성과물도 특허출원되어 권리화되고 있는 것을 알 수 있다. 이외에도 수시로 수집된 대용량의 비정형 데이터(빅데이터)는 그 전송과 관리, 저장 등에 관한 특허도 많이 등록된 것을 알 수 있다.

이와 같이 데이터 관련 특허권에 의하여 생성된 데이터는 정보재로서 상거래의 대상이 되고, 그 데이터의 속성이나 특성 등의 정보분석을 하는 데이터마이닝이 이루어지고 있고, 그 결과를 이용하여 AI학습용 데이터구조가 가공된다. 여기서 특허권자의 허락 없이 데이터를 가공하는 행위는 특정 목적의 AI학습용으로 데이터전처리가 이루어지는 것이므로 특허침해가 성립하지만, 특정 목적이나 상업적 이용을 목적으로 하지 않고 데이터의 속성이나 특성을 알기 위한 데이터마이닝을 하는 행위는 특허법 제96조의 “시험 또는 연구”에 해당하여 특허권의 효력이 제한되는 것인지가 명확하지 않은 것을 알 수 있었다.

따라서 데이터의 사용 및 그 활용을 적극 추진하여 새로운 비즈니스모델을 도출하고, 이를 이용한 새로운 비즈니스모델을 적용한 신산업 및 신제품 등을 추진하고 촉진하기 위해서는 데이터마이닝을 “시험 또는 연구”에 포함하여 특허권의 효력이 제한되는 것으로 특허법의 개정이 필요하다.

참고문헌

〈단행본(국내 및 동양)〉

- 권태복, 『제4차 산업혁명과 특허전략』, 한국지식재산연구원·세창출판사, 2019.
- 송영식·이상정·황중환, 『지적소유권법(상)』, 육법사, 2008.
- 윤선희, 『특허법』, 제6판, 법문사, 2019.
- 中山信弘·小泉直樹, 『新·注解 特許法(中卷)』, 第2版, 青林書院, 2017.
- 經濟産業省 情報經濟課, 『AI・データの利用に関する契約ガイドラインと解説』, 商事法務(NBL), 2018.

〈학술지(국내 및 동양)〉

- 권지현, “비기술적 요소가 포함된 AI발명의 진보성 판단기준”, 『법학논총』, 제48집 (2020).
- _____, “AI발명의 성립성과 진보성 판단의 조화방안”, 『홍익법학』, 제22권 제3호 (2021).
- 박충목·차상욱, “인공지능(AI)관련 발명의 지식재산권법상 보호방안: 특허법 및 영업비밀보호법을 중심으로”, 『Issue Paper』, 제2019-11호(2019).
- 김성원·최상민·이수원, “텍스트·데이터 마이닝 과정의 저작물 이용 면책 규정 신설안에 대한 소고-적대적 생성신경망에의 적용”, 『지식재산연구』, 제18권 제1호(2023).
- 김창화, “실험적 사용에 의한 특허권 제한 범위의 고찰- 미국법과의 비교를 중심으로”, 『財産法研究』, 제34권 제1호(2017).
- 박성호, “저작권법 전면 개정안 주요 내용의 비판적 고찰”, 『정보법학』, 제24권 제3호 (2020).
- 박영규, “연구 또는 시험을 하기 위한 특허발명 실시의 의미와 한계”, 『산업재산권』, 제31호(2010).
- 박진아, “데이터의 보호 및 유통 법제 정립 방안”, 『서강법률논총』, 제9권 제2호 (2020).
- 이규호, “인공지능 학습용 데이터세트 보호를 위한 특허법상 주요 쟁점 연구”, 『산업재산권』, 제64호(2020).
- 이상용, “데이터의 비계약적 이용-데이터 마이닝을 위한 저작권 제한을 중심으로”, 『강원법학』, 제65권(2021).
- 조영선, “研究・試驗을 위한 特許發明의 實施와 特許權의 效力”, 『저스티스』, 통권 제116호(2010).

차상욱, “저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점—텍스트·데이터마이닝(TDM) 면책규정”, 『경영법률』, 제32집 제1호(2021).

홍석경, “리서치툴 특허의 이용에 관한 연구”, 『지식재산연구』, 제10권 제3호(2015).

홍승기, “데이터마이닝 면책 입법 방향에 대한 의문”, 『경영법률』, 제32집 제4호(2022).

〈학술지(서양)〉

David L. Parker, “Symposium, Patent Infringement Exemptions for Life Science Research”, *HOUS. J. INT’L. L.*, Vol.16 (1994).

Shlomit Yanisky Ravid & Xiaoqiong (Jackie) Liu, “When artificial intelligence systems produce inventions: An alternative model for patent law at the 3A era”, *Cardozo L. Rev.*, Vol.39 (2018).

Jordan Karp, “Note, Experimental Use as Patent Infringement: The Impropriety of a Broad Exemption”, *YALE L.J.*, Vol.100 (1991).

Metzger, Patentverletzung durch Forschung, GRUR Int. 1967, S. 124, 126.

Richard E. Bee, “Experimental Use as an Act of Patent Infringement”, *J. PAT OFF. SOC’Y*, Vol.34 (1957).

〈연구보고서(국내 및 동양)〉

한국데이터산업진흥원, “2022 데이터산업백서”, 통권 제25호, 한국데이터산업진흥원, 2022.

〈기타 자료(국내 및 동양)〉

특허청, 『기술분야별 심사실무가이드』, 특허청, 2023.

Data Mining and Restrictions on the Effect of Patent Rights

Kwon, Chihyun

Data obtained by the implementation of data generation and collection method patents (data patents) is an information material that is subject to commerce, and is processed into a data structure and traded. In the case where a third party performs data mining without the permission of the patent holder, the issue arises in relation to whether such act violates data patent rights.

Article 96 of the Patent Act stipulates that the effect of a patent right is limited when a patented invention is implemented for testing or research activities, but it is not clear whether data mining corresponds to testing or research activity, and there is another issue that new industries applying AI with data cannot be promoted more actively.

To solve this issue, it is necessary to review a revised plan that requires data mining to be included in testing or research activities. In addition, in order to strengthen the protection of data patent holders, Article 129 needs to be revised so that the estimation regulations of production methods can be applied by considering data generated by the implementation of generation and collection method patents as objects or media. It is also necessary to further consider revising Article 2, Paragraph 3 so

that data is included in objects that fall within the scope of the invention of the object.

Keywords

Data, data generation, data collection, data structure, test or study, effect of patent rights, restriction on patent rights