

# 디지털 혁신과 데이터법

김현경 교수

서울과학기술대학교 IT정책전문대학원

hkyungkim@seoultech.ac.kr



## 데이터법의 의미와 범위

디지털 생태계에서 데이터가 차지하는 비중과 역할의 중요성에 대하여는 이미 많은 연구가 이루어진 바, 본 챕터에서는 그러한 원론적 논의는 생략하고 '데이터법'에 대한 논의에 집중하고자 한다. '데이터법'은 '데이터'를 규율하는 법을 총칭하여 '데이터법'으로 칭할 뿐, '민법'이나 '형법'처럼 단일법으로 규율되어 있는 것이 아니다. 데이터의 유형에 따라서 인간의 창작적 노력이 투입된 데이터를 지적재산(지식재산권법)으로, 정보주체를 식별하는 인격적 권리가 투영된 데이터를 개인정보(개인정보 보호법)로 보호한다. 또한 일정한 물적·인적 투자가 투입된 데이터베이스(저작권법)를 보호하기도 한다. 이처럼 법으로 보호가 규정된 데이터를 제외하면 원칙적으로 그 외 데이터의 이용은 자유다. 이러한 데이터법의 방향은 전 세계가 거의 유사하다. 다만 그 보호의 정도가 다를 뿐이다.

우리나라를 기준으로 데이터 유형에 따른 데이터법 현황은 다음 표와 같다.

디지털 혁신 환경에서 활용 가치가 높은 데이터는 단연코 개인정보다. 또한 최근 인공지능의 학습용 데이터로 수많은 저작물이 이용되고 있는 바, 이러한 경우 저작권의 제한을 인정하는 입법이 국내외에서 추진되고 있다. 개인정보 역시 2020년 개정으로 가명처리를 통해 개인정보의 활용을 꾀할 수 있게 되었다. 데이터의 이동은 물리적 객체와는 달리 탈국경성을 속성으로 하는 바, 데이터와 관련된 입법은 대부분 국제적 통일성을 향하며 국제규범으로부터 긴밀한 영향을 받는다. 이하에서는 디지털 혁신에 대응하는 유럽과 우리의 데이터법 동향을 살펴보고 추후 입법과제를 전망해 본다.

[표] 데이터유형과 데이터법 개관<sup>1)</sup>

데이터 유형	데이터법	핵심내용
지식재산	- 산업재산: 특허법(발명), 실용실안법(고안), 디자인보호법(디자인), 상표법(상표) - 저작권법(저작물)	- 보호 중심 - 주로 재산권적 보호(*저작권은 인격권도 보호) - 독점배타적 권리 부여
데이터베이스	- 저작권법 - 부정경쟁방지법	- 보호 중심 - 독점배타권 또는 금지청구권
개인정보	- 개인정보 보호법(개인정보) - 신용정보법(개인신용정보) - 위치정보법(개인위치정보)	- 보호 중심 - 인격권적 보호 - 정보주체의 자기결정권
공공데이터	- 공공데이터법(공공데이터)	- 활용 중심 - 공공데이터의 민간 활용 촉진
그 밖의 데이터	- 부정경쟁방지법 (일정한 요건의 데이터, 영업비밀)	- 보호 중심 - 금지청구권

1) 부정경쟁방지 및 영업비밀보호에 관한 법률(약칭: 부정경쟁방지법)

## 유럽의 움직임: '보호'의 자물쇠를 열며

유럽은 지식재산, 개인정보, 데이터베이스 등에 대하여 강력한 보호 중심의 입법방식을 채택하였다. 특히 개인정보에 대하여는 유럽 일반 개인정보보호법(GDPR)<sup>2)</sup>이라는 강력한 보호 중심의 규범을 통해 국제적 표준화를 꾀하고 있다. 뿐만 아니라 이미 25년 전인 1996년 데이터베이스에 대한 투자를 독려하여 데이터 시장을 개척하기 위해 창작적 요소가 없는, 즉 저작권으로 보호되지 않는 데이터베이스에 대한 독점배타적 권리를 인정하는 지침을 채택한 바 있다.<sup>3)</sup> 그러나 오히려 이러한 권리를 부여하지 않은 미국의 데이터베이스 생산량이 유럽에 비해 상당하며, 특히 유럽에서도 이러한 보호 중심의 데이터베이스권 도입 후의 관련 경제적 효과에 별반 탁월한 차이가 발견되지 않았다.

이러한 상황에서 EU는 4차 산업의 핵심 산업인 데이터 산업에서 EU의 주도권 확보를 위해 역내 데이터 단일시장을 형성하고, 합법적이고 원활한 데이터 공유 환경을 조성할 수 있도록 2020년 EU 데이터 전략(European Strategy for Data)을 수립하였다. 이 전략 실행을 위한 법안을 단계적으로 추진하는데, 첫 번째 단계로 추진된 것이 2021년 11월 발표된 데이터 거버넌스법(DGA, Data Governance Act)이다. 이 법은 EU 내 데이터 주권을 수립하고, 데이터가 자유롭고 안전하게 이동할 수 있도록 법적 프레임워크를 마련하는 것이다. 두 번째 단계로 2022년 2월 23일 발표된 데이터법(Data Act) 초안이다.<sup>4)</sup> 이 법안은 거버넌스법의 법적 토대를 바탕으로 구체적 인 데이터 접근 및 제공 주체와 공유 조건을 규정하고 있다.<sup>5)</sup> Data Act의 목표는 IoT 기기에서 생성되는 데이터에 대하여

기존 데이터 시장의 약자로 여겨지는 개인과 중소기업의 접근성을 강화해 공유를 활성화하는 것이다. Data Act은 누가 어떻게 데이터에 접근할 수 있는지와 어떻게 가치를 창출할지에 대한 조건을 규정하고 있다. 또한 데이터 공유의 안전성과 합법성을 확보하기 위해 보안 규정을 강조하고 있다. 즉 Data Act은 데이터 접근과 이동을 보장하기 위한 특단의 조치로 IoT 기기 사용자에게 데이터에 접근 및 제공할 수 있는 새로운 권리를 인정하고 있으며 그 권리실행을 위해 IoT 제조업자에 해당하는 데이터베이스 권리자에게 데이터 이동 의무를 부여하고 있다.

Data Act은 IoT 제품을 사용하고, 산업 데이터 생산에 기여하는 개인 사용자의 접근 권한을 강화하고 있다. 따라서 사용자는 자신이 생산에 기여한 데이터에 언제든지 접근할 수 있으며, IoT 제조업자로 하여금 사용자가 지정한 제3자에게 데이터를 전송하도록 요청할 수 있다. 여기서 제3자는 제품의 유지, 보수 및 컨설팅과 같은 애프터 마켓서비스를 제공하는 업체를 말한다. 주로 중소기업 대상으로 하며 이미 시장에서 지배력을 행사하고 있는 게이트키퍼(Gatekeeper) 기업<sup>6)</sup>은 이와 같은 무상 제공 대상에서 제외된다.

이러한 사용자의 데이터 접근 및 이동권은 IoT 제조업자 등의 데이터베이스권과 상충되기 때문에 Data Act은 기계 생성 데이터의 경우 우선 데이터베이스 지침이 인정하는 데이터베이스권의 예외를 인정함으로써 자유로운 재사용을 허용하고 있다.<sup>7)</sup> 이처럼 유럽은 데이터에 대한 보호 중심의 법체계를 고수하였으나, 그러한 시스템하에서 데이터 경제의 획기적 발전을 기대하기 어렵다는 것을 인식하고 적극적으로 데이터를 활용하기 위한 입법적 조치를 강구하고 있다.

2) General Data Protection Regulation, 2018년 5월 25일부터 시행되고 있는 유럽연합(EU)의 일반 개인정보보호법으로 EU 내 사업장이 없더라도 EU를 대상으로 사업을 하는 경우 적용대상이 될 수 있다.

3) Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of database.

4) 원 제목은 "데이터에 대한 공정한 접근 및 사용에 관한 통일된 규칙에 대한 규정(Regulation on harmonised rules on fair access to and use of data)"이다.

5) 마지막으로 EU는 산업별로 역내 공동 데이터 공간을 개발해 본격적으로 데이터 공유를 추진할 계획이다. 현재 개발이 예정된 10개 핵심 분야는 건강, 농업, 제조, 에너지, 운송, 금융, 공공행정, 기술, 연구 데이터, 그린딜이며 이 중 가장 먼저 건강 분야의 공동 데이터 공간에 대한 입법이 준비 중이고, 곧 자동차에서 발생하는 산업 데이터의 공유 규칙을 다루는 운송 분야 데이터 공간도 발표될 예정이다.

6) EU의 디지털 시장법(DMA)에서는 게이트키퍼 기업을 연간 매출액 75억 유로 이상, 시가총액 750억 유로 이상, 월간 사용자 4,500만 명 이상, 연간 비즈니스 사용자 1만 명 이상 인 플랫폼 서비스로 규정(EU 의회 DMA 합의 발표(3.24.))

7) <https://www.euractiv.com/section/data-protection/news/leak-draft-impact-assessments-sheds-some-light-on-upcoming-data-act/> (2022.7.24. 최종 확인)

## 우리의 대응: 데이터 입법의 홍수

2022년 기준, 데이터 관련법들은 앞에서 기술한 법률 외에도 과학기술정보통신부 소관의 데이터산업법<sup>8)</sup>과 산업통상자원부 소관의 산업디지털 전환법<sup>9)</sup> 특허청 소관의 부정경쟁방지법<sup>10)</sup>이 입법되어 시행되고 있다. 데이터산업법과 산업디지털 전환법이 데이터와 관련된 산업의 진흥과 시책을 추진하는 것을 주된 목적으로 하는 것이라면, 부정경쟁방지법은 데이터에 대한 부정사용을 금지함으로써 데이터 관련 권익의 보장을 주된 내용으로 하는 것이다. 뿐만 아니라 중소기업 스마트제조혁신 지원에 관한 법률안<sup>11)</sup> 국가데이터위원회의 설치 및 운영에 관한 법률안,<sup>12)</sup> 데이터재산권의 보호 등에 관한 법률안,<sup>13)</sup> 문화데이터 조성 및 활용 촉진에 관한 법률안<sup>14)</sup> 등이 국회에서 논의되고 있다. 다만 데이터산업법을 비롯한 산업디지털 전환법 등은 개인정보, 저작권, 공공데이터 등에 대하여는 기존의 개인정보 보호법, 저작권법 등이 적용되도록 규정하고 있는 바, 이러한 법들은 앞서 언급한 바와 같이 거의 대부분 유럽의 입법례를 따르고 있다. 데이터베이스에 대하여는 유럽의 1996년 지침을 그대로 벤치마킹하여 2003년 저작권법 개정을 통해 독점배타적 권리를 인정한 바 있으며,<sup>15)</sup> 2020년 개인정보 보호법의 개정 역시 가명정보의 처리 등 유럽의 GDPR을 상당 부분 반영한 것으로 보인다.

실상 데이터 활용을 위한 다수의 입법을 각 부처가 경쟁적으로 추진하였지만 실제 산업계의 활용이 요구되는 개인정보, 저작물 등에 대하여는 여전히 강력한 보호 중심의 법체계를 유지하고 있는 것이다. 특히 데이터산업법과 산업디지털 전환법은 표준계약서, 사업자 규제, 데이터 품질관리, 데이터 플



랫폼 등 상당한 시책이 양법에서 중복하여 중첩적으로 시행될 예정이라 오히려 수범자 혼란을 야기할 수 있다. 뿐만 아니라 산업디지털 전환법은 산업데이터에 대한 사용·수익할 '권리'를 인정하고 있는데(제9조 제1항부터 제3항) 이러한 권리가 새로운 독점 배타적 권리의 창설인지, 단순히 채권적 권리인지 불분명하므로 실제 법 집행 및 수범자 혼란이 우려된다. 또한 산업데이터가 저작권법상 데이터베이스에 해당될 경우 저작권법은 데이터베이스를 공동으로 제작한 경우 공동의 합의에 의해서만 사용할 수 있으나, 산업디지털 전환법은 각자 사용·수익할 권리를 인정한다고 하고 있으므로 양자의 충돌도 우려된다.

이처럼 데이터 입법의 홍수라 불릴 만큼 데이터 활용을 위한 다수의 입법을 추진 중이나, 정작 중요한 쟁점인 인공지능 학습용 데이터로서 저작물이나 개인정보의 활용, 공개된 개인정보의 활용, 데이터의 크롤링 등의 법적 해결에 있어서 여전히 미온적이다.

8) 데이터 산업진흥 및 이용촉진에 관한 기본법, 법률 제18475호, 2021년 10월 19일 제정, 2022년 4월 20일 시행

9) 산업 디지털 전환 촉진법, 법률 제18692호, 2022년 1월 4일 제정, 2022년 7월 7일 시행

10) 법률 제18548호, 2021년 12월 7일 일부개정, 2022년 4월 20일 시행

11) 송갑석 의원 대표발의, 2020.8.26. 제안, 의안번호 2103284

12) 박성중 의원 대표발의, 2021.4.13. 제안, 의안번호 2109446

13) 박성중 의원 대표발의, 2021.4.13. 제안, 의안번호 2109453

14) 도중환 의원 대표발의, 2021.11.23. 제안, 의안번호 2113494

15) '데이터베이스'는 소재를 체계적으로 배열 또는 구성한 편집물로서 개별적으로 그 소재에 접근하거나 그 소재를 검색할 수 있도록 한 것이고, '데이터베이스 제작자'는 데이터베이스의 제작 또는 그 소재의 갱신·검증 또는 보충에 인적 또는 물적으로 상당한 투자를 한 자이다(저작권법 제2조). 데이터베이스 제작자는 자신의 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송할 권리를 가진다(저작권법 제93조 제1항).

## 데이터법의 전망과 과제



우리가 주로 벤치마킹한 유럽은 데이터에 대한 강력한 보호 중심의 입법을 채택하였지만, 그 경제적 효과가 증명되지 못한 채 앞서 검토한 바와 같이 스스로 수정을 시도하고 있다. 오히려 이러한 보호를 취하지 않은 미국 등이 데이터 경제의 최강국이 되었다는 점을 볼 때 데이터 경제에 있어서 보호 중심의 입법은 한계를 지닌다고 볼 수 있다.

그럼에도 불구하고 각 국가별로 처한 상황과 고유한 여건은 다르겠지만 디지털 혁신을 추구함에 있어서 데이터의 활용을 위한 다음과 같은 과제는 반드시 해결되어야 한다는 데에는 공감할 것이다.

첫째, 인공지능 학습용 데이터의 증진이다. 인공지능 학습은 데이터전처리(Pre-Processing)을 거친 데이터의 전부 또는 다수를 학습데이터로 입력하는 경우가 많으므로 데이터에 대한 적법한 이용 권한이 없다면 대규모의 저작권 침해가 발생할 수 있다. EU, 일본, 영국 등은 이를 위한 저작물의 데이터 처리 관련 면책요건을 입법하였고, 미국은 판례법에 근거한 공정이용 법리<sup>16)</sup>에 따라 이를 허용하는 논의를 하고 있다. 우리나라도 2021년의 저작권법 전부 개정안<sup>17)</sup>에서 TDM(Text and Data Mining)을 위한 예외를 규정하여 인공지능 알고리즘 학습을 위한 저작권 제한 규정을 마련하고자 하고 있다. 다만 해당 저작물에 적법하게 접근할 수 있는 경우에 한정하여 허용하고 있으므로 실무적으로 이루어지는 웹사이트 등에 대한 데이터 크롤링, 스크레이핑이 적법한지는 입법이 되더라도 여전히 문제될 수 있다.

둘째, 공공데이터의 적극적 개방과 활용이다. 우리나라는 「공공데이터의 제공 및 이용 활성화에 관한 법률」을 통해 공공데이터의 적극적 개방과 민간 활용을 추구하고 있다. 그러나 본법의 적용대상에서 저작권법상의 데이터베이스는 제외하고 있다. 더욱이 데이터베이스권의 보호의 대상이 되는 데이터베이스 관련하여 법원은 로마켓(lawmarket.co.kr) 사건<sup>18)</sup>에서 “사건검색 서비스를 제공하기 위하여 법원은 인적·물적으로 상당한 투자를 하여 소송사건 관련 정보 데이터베이스를 제작하였을 것이므로 법원은 그에 대하여 데이터베이스 제작자로서의 권리를 가지고 있다”고 하여 법원도 즉 공공도 데

16) Authors Guild v. Google 사건(Authors Guild v. Google Inc., 804 F.3d 202 (2d Cir. 2015)): 구글이 특정 검색 키워드가 텍스트에서 몇 번 발견되는지 등의 저작물 텍스트에 관한 사실적 정보(Factual Information)를 대중에게 제공하기 위해 저작물을 복제하는 것은 공정이용에 해당한다고 판시하여 저작권 침해를 부정하였다.

A.V. ex. rel. Vanderhye v. iParadigms 사건(A.V. ex. rel. Vanderhye v. iParadigms, L.L.C., 562 F.3d 630 (4th Cir. 2009)): 양 저작물의 텍스트적 유사성을 비교하는 것은 저작물의 창작적인 요소와 관계가 없다는 이유를 들어 iParadigms가 표절탐지도구에 사용하기 위해 학생들의 논문을 복제한 행위는 공정이용에 해당한다고 하여 저작권 침해를 부정하였다.

17) 도종환 의원 대표발의, 저작권법 전부개정안, 2021.1.15. 의안번호 7440 제43조(정보분석을 위한 복제·전송) ① 컴퓨터를 이용한 자동화 분석기술을 통해 다수의 저작물을 포함한 대량의 정보를 분석(규칙, 구조, 경향, 상관관계 등의 정보를 추출하는 것)하여 추가적인 정보 또는 가치를 생성하기 위한 것으로 저작물에 표현된 사상이나 감정을 향유하지 아니하는 경우에는 필요한 한도 안에서 저작물을 복제·전송할 수 있다. 다만 해당 저작물에 적법하게 접근할 수 있는 경우에 한정한다.

18) 서울고등법원 2008.4.16. 선고 2007나74937 판결 [정보계시금지등](원심: 서울중앙지방법원 2007.7.6. 선고 2006가합22413 판결 [정보계시금지등] [각공2007.8.10. (48),1590]: 피고는 “lawmarket.co.kr”이라고 하는 인터넷 홈페이지 운영자로 법원의 사건검색 데이터베이스를 이용하여 원고들(변호사들)에 대한 정보를 모아서 검색 결과를 써 제공하였다. 본 사안에서 자기정보통제권, 국민의 알권리의 대상, 개인정보 해당성 등 복합적 쟁점을 다루었으나, 본 챗터에서는 관련된 데이터베이스권에 대한 사안만 다룬다.

이터베이스권의 주체가 될 수 있음을 밝혔다. 즉 공공데이터베이스는 '자유이용'이 아니라 '보호'의 대상인 것이다. 저작권법상 데이터베이스 제작자의 권리제한(제94조)에서 공공저작물의 자유이용(제24조의2) 규정을 준용하고 있지 않으므로 엄격히 공공데이터베이스에 대하여 데이터베이스 제작자의 권리 제한이 인정되지 않는다고 볼 수 있다. 그러나 공공저작물과 마찬가지로 공공데이터베이스 역시 국민의 세금으로 제작되며 비배타성 등으로 특정인의 이용이 다른 이의 이용을 배제하지 않아 이용에 특혜의 문제가 발생하지 않는다. 그럼에도 불구하고 공공저작물과 달리 취급하는 것은 적절하지 않다. 또한 이미 「공공데이터의 제공 및 이용 활성화에 관한 법률」에서 공공기관이 보유·관리하는 공공데이터의 민간 제공을 원칙으로 규정하고 있는 바,<sup>19)</sup> 추후 양법의 충돌 가능성도 고려할 필요가 있다.

셋째, 개인정보의 합리적 활용이다. 일정한 요건하에 공개된 개인정보의 활용을 허용하는 방안, 인공지능 학습용 데이터로 개인정보의 활용을 허용하는 방안 등이 모색될 필요가 있다. 개인정보를 보호하는 이유는 정보주체의 인격적 법익에 대한 침해를 막기 위함이다. 인공지능 학습용 데이터로 활용되는 경우 그 인격적 법익의 침해가능성이 상당히 낮고 이미 모두에게 공개된 개인정보의 경우도 그러하다. 마이데이터를 비롯한 개인정보의 활용을 통한 정보주체의 수익추구 비즈니스도 제도화되고 있는 만큼 개인정보의 합리적 활용은 데이터법의 가장 큰 과제일 것이다.



19) 제17조(제공대상 공공데이터의 범위) ① 공공기관의 장은 해당 공공기관이 보유·관리하는 공공데이터를 국민에게 제공하여야 한다. 다만 다음 각 호의 어느 하나에 해당하는 정보를 포함하고 있는 경우에는 그러하지 아니한다.

1. 「공공기관의 정보공개에 관한 법률」 제9조에 따른 비공개대상정보

2. 「저작권법」 및 그 밖의 다른 법령에서 보호하고 있는 제3자의 권리가 포함된 것으로 해당 법령에 따른 정당한 이용허락을 받지 아니한 정보