

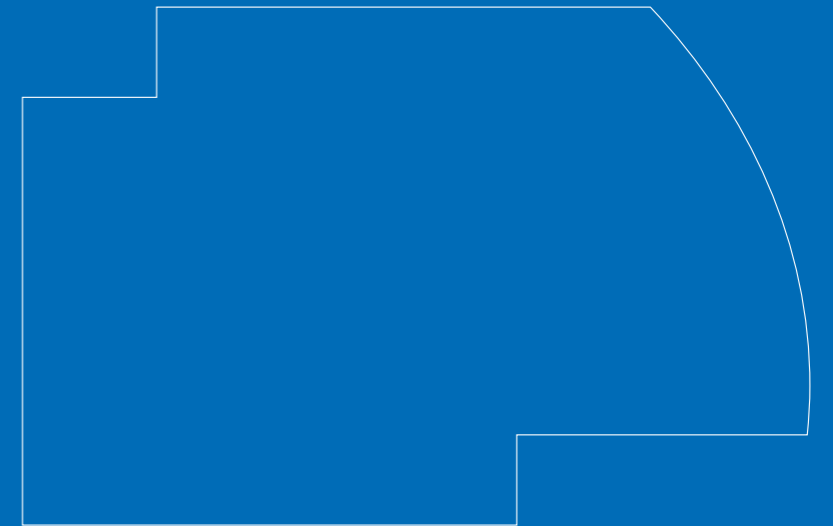
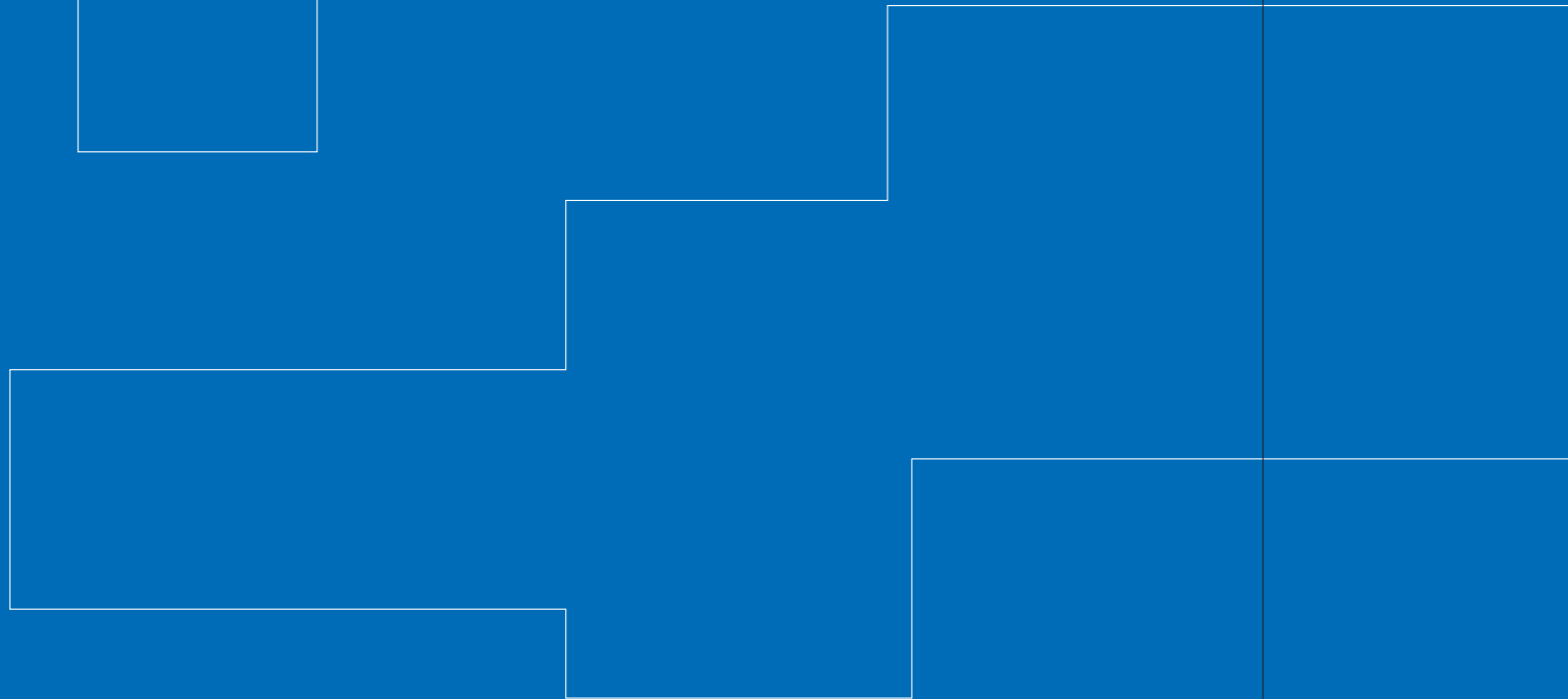
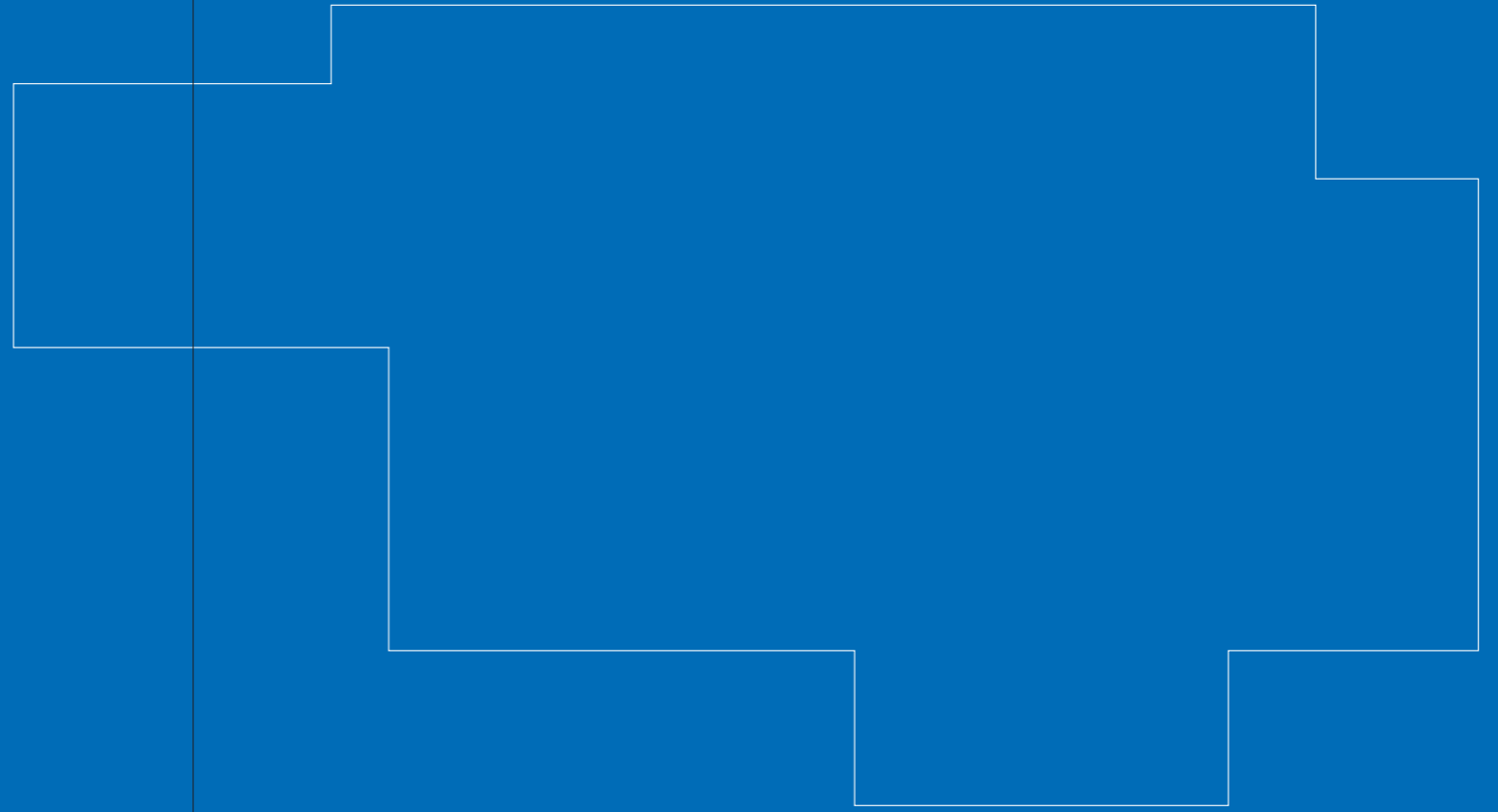
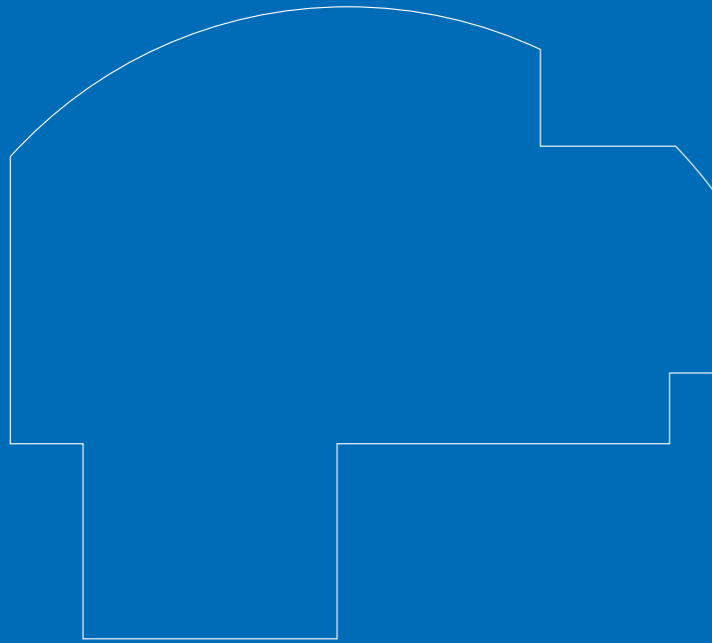
이슈페이퍼 2020-1

데이터 3법 시대의 과제:

- 가명처리
- 연구목적 활용
- 차등적 프라이버시
- 데이터 거래

데이터거래의 법적 쟁점 및 데이터거래 가이드라인

이동진 서울대학교 법학전문대학원 교수



데이터가 오늘날 가치의 주요원천이라는 점에는 더 이상 의심이 없다. 그런데 데이터는 수집이 어려운 반면 일단 수집되면 공유·재사용은 쉽다는 특성이 있다. 그리하여 데이터의 공유와 재사용을 가능하게 하는 자발적 메커니즘으로서 데이터거래에 관심이 모이고 있다.

데이터거래는 거래대상인 데이터의 법적 성격이 분명하지 아니하여 그 법적 구성에 불확실한 점이 있고, 데이터경제 자체가 부분적으로 기존의 시장과 다른 특성을 보여 추가적인 법적 쟁점이 생길 가능성도 있다. 데이터거래를 둘러싼 법적 위험을 인식하고 거래여부결정과 계약내용형성과정에서 이에 적절히 대응하는 것이 그 활성화에 필수적이다. 데이터소유권(data ownership)을 둘러싼 각국의 논의는 대체로 데이터소유권의 인정에 힘쓰기보다 우선은 표준계약과 가이드라인의 제시를 통하여 데이터거래계약을 지원할 필요가 있다고 한다.

현행법상 데이터거래와 관련하여 우선 구별하여야 할 점은 거래대상이 개인정보인 데이터인지 아니면 개인정보가 아닌 데이터인지이다. 개인정보인 데이터의 경우 거래를 하려면 정보주체의 동의를 받거나 비식별화·익명화를 거쳐야 한다. 주의할 점은 비식별화·익명화를 거친다 하더라도 대개 데이터 수령자로 하여금 계속 기술적·관리적 조치를 취하게 할 필요가 있다는 사실이다. 영업비밀로 관리되어온 데이터는 영업비밀 기타 부정경쟁방지법과 불법행위법에 의하여 보호된다. 그러나 이때에도 그러한 보호를 유지하려면 데이터 수령자로 하여금 비밀로 관리하게 할 필요가 있다. 그밖에 이른바 파생데이터의 귀속에 대하여는 현행법의 규율이 불분명하다. 계약적 규율이 필요한 까닭이다.

데이터거래와 관련하여서는 유럽연합이 제시한 오픈 데이터 접근, 데이터 시장에서의 금전화 및 폐쇄된 플랫폼에서의 데이터 교환의 구분과 일본 경제산업성 가이드라인이 제시한 데이터제공형, 데이터창출형 및 데이터공용형계약의 구분이 참고가 된다. 두 구분은 상당부분 중첩되는데, 그중 각각의 마지막인 플랫폼형은 플랫폼규약의 설계라는 점에서 개별적인 B2B계약과 구별된다. 이 보고서에서는 우선 개별적인 B2B계약 유형에 집중하였다.

데이터거래, 데이터계약에서 먼저 고려할 점은 데이터의 처분권한 유무와 데이터공유의 경쟁에 대한 영향 및 경쟁법상의 제한이다. 데이터공유는 종종 제공자의 경쟁상 우위를 손상시킬 수 있다. 그러나 다른 한편 필수적 데이터의 제공거부는 경쟁법 위반이 되거나 적어도 비난의 대상이 될 수 있다. 다음 고려할 점은 데이터의 경제적 가치 평가인데 아직까지는 난점(難點)이 많다. 창출된 가치에 참여하는 모델을 함께 고려할 필요가 있다. 데이터계약의 유형으로는 양도와 이용허락이 대표적인데, 전자에는 매매법을, 후자에는 제공자의 동시이용여부에 따라 임대차법이나 지식재산권이용허락의 법리를 적용할 수 있다. 데이터창출형은 도급이다. 특히 데이터양도의 경우 실제로는 데이터삭제를 확인할 절차를 규율할 필요가 있다. 그밖에 데이터거래에서는 포맷과 제공방법,

데이터보안을 위한 사전·사후조치와 협력, 데이터에 관한 지식재산권 등 제3자 권리에 관한 담보책임과 이용허락의 범위 특정, 데이터 수령자 측의 이용권한의 객관적·주관적 범위 한정, 개인정보인 데이터의 경우 개인정보 보호법상의 제공 요건의 충족 및 철회를 대비한 보완조치에 관하여 정하여야 한다. 가명화된 데이터의 사용과 데이터 결합의 경우 개인정보 보호법상의 요건 충족 여부를 신중히 검토하여야 한다.

데이터거래, 특히 개인정보이거나 개인정보에서 유래한 데이터의 거래는, 법적 요건이 아니라 하더라도, 투명성과 가치공유에 대한 배려가 필요하다는 점을 염두에 둘 필요가 있다.

1. 들어가는 말

2019년 현재를 포함하여 지난 몇 년간 시가총액 기준 전 세계에 가장 큰 기업 1~5위에는 마이크로소프트(Microsoft), 아마존(amazon), 애플(apple), 구글(google, 회사명은 alphabet), 페이스북(facebook)이 포함되어 있었다. 모두 IT기업이고, 상당수는 데이터기업이기도 하다. 특히 그중 구글과 페이스북은 주로 데이터를 활용하는 플랫폼(platform)기업이다.

근래 전 세계적으로 데이터경제(data economy)가 중요해졌다는 점에는 이론(異論)이 없다. 데이터 수집·저장 및 처리기법과 플랫폼 영업모델의 발전으로 데이터를 활용한 가치창출이 비약적으로 증가하였다. 이제 누구도 미래의 성패가 데이터에 달려있음을 의심하기 어렵다.

그런데 데이터에는 모으면 모을수록 그 가치가 비약적으로 증가하지만 대량으로 모으는 데 상당한 비용과 시간이 들고, 반면 일단 모으면 복제에는 비용이 거의 들지 아니하며 품질도 훼손되지 아니한다는 특징이 있다. 일단 모은 데이터를 공유(data sharing) 내지 재사용(re-use)하여 데이터의 활용을 극대화할 필요가 제기되는 까닭이다. 이는 데이터거래를 전제한다. 데이터거래는 이미 행해지고 있고, 앞으로 더욱 확대될 것이며, 또 더욱 확대할 필요가 있다. 여러 나라에서 근래 데이터 공유 및 재사용을 촉진하기 위한 방안에 고심하고 있다.⁰¹

이 글에서는 데이터거래가 어떤 법적 구조를 취하고 있고 그 쟁점은 무엇인지를 살펴보고자 한다. 그리하여 현재 데이터거래를 하고자 하는 경우 어떤 문제를 검토하여야 하는지를 밝힘과 동시에 장래 데이터거래를 촉진하기 위하여 취할 수 있는 조치에 관하여도 시사점을 도출하고자 한다.

2. 데이터의 구분과 데이터에 대한 권리

가. 개인정보인 데이터와 비개인정보인 데이터

1) 데이터

데이터(data)란 무엇인가. ‘데이터’는 ‘정보(information)’와 혼용되고 있는 개념으로, 어떤 사실을 정신적 작용을 통하여 인식, 표출한 것을 가리킨다.⁰² 오늘날에는 정보를 내용/의미론적 정보(content layer, semantische

⁰¹ 대표적으로 미국의 Sharing Data While Protecting Privacy (M-11-02 of November 3, 2010), Open Data Policy-Managing Information as an Asset (M-13-13 of May 9, 2013) 및 Next Steps in the Evidence and Innovation Agenda (M-13-17 of July 26, 2013), 유럽연합(EU)의 The Communication “Towards a data-driven economy” [COM(2014) 442 final of 2 July 2014] 및 Digital Single Market strategy [COM(2015) 192 final of 6 May 2015] 및 일본의 AI·데이터의利用に する契約ガイドライン, 2018을 들 수 있다. 우리나라도 2018. 6. 관계부처 합동으로 ‘데이터 산업 활성화 전략’이 4차산업혁명위원회 의결안건으로 제출되어 의결된 바 있는데, 그 내용 중 중요한 한 부분이 데이터의 유통이었다.

⁰² 오병철, 디지털정보계약법, 2005, 4-5면.

Information), 표시·기호/구문론적 정보(code layer, syntaktische Information) 및 물리적·유형적 실현물/구조적 정보(physical layer, strukturelle Information)로 구분하는 경우가 많다.⁰³ 예컨대 책이나 그림, 건축물이 물리적·유형적 실현, 즉 구조적 정보라면 그곳에 담긴 문자열, 음률, 화상, 구조설계 등이 표시·기호, 즉 구문론적 정보에 해당하고 이들의 감상·해석하여 얻을 수 있는 의미, 심상(心想)이 내용, 즉 의미론적 정보이다. 데이터는 그중 인간이 인식하여 표출한 기록형태, 즉 구문론적 정보를 가리킨다. 오늘날에는 발전된 데이터처리기술로 인하여 인간이 직접 인식하여 코딩(coding)할 필요 없이 기계가 코딩하는 경우가 증가하게 되었는데, 그러한 경우도 결국 직·간접적으로 인간의 인식이 가능한 형태로 정보를 추출, 정리하였다는 점에서 데이터라고 할 수 있다. 이처럼 정보의 여러 층위를 구별하는 것은 단일한 정보가 관계되는 것처럼 보여도 그 서로 다른 측면, 가령 구조적 측면과 구문론적 측면, 의미론적 측면에 대하여 각각 별도의 권리가 설정되고 별개의 법률관계가 전개될 가능성이 있기 때문이다. 그러한 경우 데이터거래는 그중 코드를, 그리고 원칙적으로는 코드만을 대상으로 하는 것이다.

나아가 적어도 실무적으로 데이터거래에서 데이터는 사실상 디지털로 코딩된 경우에만 문제된다.⁰⁴ 컴퓨터 등 전자기기로 읽어 들이고 디코딩(decoding)할 수 있는 데이터만이 품질손상 없는 무제한적 복제가능성과 데이터 처리기법을 통한 가치창출 가능성을 갖고 있고, 오늘날 논의될 필요성이 있는 것이다.⁰⁵

끝으로 데이터는 다양한 정보 중 정보처리를 위한 재료(raw material)로서의 성질을 가진 것만을 가리킨다. 가령 그 자체 독립적 향유의 대상이 되는 글, 화상, 음향이나 구동 가능한 컴퓨터프로그램은 디지털콘텐츠(digital contents)라고 하고,⁰⁶ 데이터와는 구별하는 것이 보통이다.⁰⁷ 이는 데이터거래는 정보가 에너지, 물질과 함께 기본적 자원 내지 재화로 기능하게 됨에 따라 문제된 것이고 그러한

⁰³ Lessig, The Future of Ideas: The Fate of the Commons in a Connected World, 2002, p. 23; Zech, Information als Schutzgegenstand, 2012, S. 37 ff.

⁰⁴ 이 점을 지적하는 것으로, 福岡眞之介·松村英寿, データ法律と契約, 2018, 3-5頁.

⁰⁵ 이 점에서는 미국의 모델법인 Uniform Computer Information Transactions Act § 102 (10)의 “computer information”의 정의도 참조. 오병철(주 2), 8면 이하는 이를 포괄하여 ‘디지털정보’라는 용어를 제안하고 있다. 한편 정보화촉진기본법 제3조 제1호는 “정보”를 “특정 목적을 위하여 광(光) 또는 전자적 방식으로 처리되어 부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식”으로 정의하고 있는데, 대체로 본문의 규정과 일치한다고 보인다.

⁰⁶ 이른바 Digital Contents Directive (EU) 2019/770.

⁰⁷ 미국의 Uniform Computer Information Transactions Act § 102 (35)는 “정보”를 “data, text, images, sounds, mask works, or computer programs”로 규정한다.

관점의 정보를 그 대상으로 하기 때문이다.⁰⁸ 하나의 정보라도 맥락에 따라 디지털콘텐츠가 되기도 하고 데이터가 되기도 한다. 음악을 듣거나 게임을 하는 경우 음원, 프로그램은 디지털콘텐츠이지만, 여러 음원을 분석하기 위하여 또는 게임의 소스코드나 규칙 등을 비교분석하기 위하여 파일을 모으는 경우는 데이터이다.

2) 구별기준

데이터의 구별기준 중 법적인 관점에서 가장 중요한 것은 개인정보인 데이터와 비개인정보인 데이터의 구별이다. 이는 헌법재판소와 대법원이 헌법상 기본권이자 사인(私人) 사이에서도 효력이 있는 권리로 개인정보자기결정권을 인정하고 있고,⁰⁹ 특히 개인정보 보호법을 비롯한 일련의 개인정보 관련 법령(신용정보의 이용 및 보호에 관한 법률, 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 위치정보의 보호 및 이용 등에 관한 법률 등)이 개인정보의 처리에 대하여 별도의 규제를 가하고 있기 때문이다. 개인정보 보호법 등이 '정보'라는 표현을 쓰지만 이는 실제로는 데이터를 가리킨다. 개인정보 보호법에서 사용하는 표현은 나라마다 다른데, 미국의 경우 '정보'(personally identifiable information; PII)라는 표현이 선호되나 유럽에서는 '데이터'(data, Daten, données)라는 표현이 선호되고, 둘은 같은 층위를 가리킨다. 우리 법의 경우 유럽 개인정보 보호법의 영향을 받았음에도 개인정보라는 표현을 쓰고 있는데, 그 표현에도 불구하고 구문론적인 층위, 즉 데이터만을 대상으로 한다는 데 별 의문이 없다.¹⁰

개인정보 보호법 제2조 제1호는 "개인정보"를 "살아있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보(해당 정보만으로는 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것을 포함한다)"로 규정한다. 반면 비개인정보는 개인정보가 아닌 일체의 정보 내지 데이터를 말한다. 즉 비개인정보는 개인정보의 잔여개념이다. 양자의 구별에 있어 핵심은 '개인을 알아볼 수 있는지', 즉 식별가능성 유무이다. 가령 일련번호 011137번이 폐암말기로 일정한 생활습관과 환경을 갖고 있었다는 내용의 데이터가 있다면, 이 데이터는 폐암의 인자를 연구하는 데 도움이 되는 정보를 포함하나, 그로부터 또는 몇 가지 정보를 더하여 적절한

데이터 처리기법을 활용하면 일련번호 011137번이 서울 강남구에 사는 ㅁ이라는 사실을 알아낼 수 있는 경우가 있다. 위 개념규정은 이러한 경우 위 데이터 전체가 개인정보가 된다는 뜻이다. 개인정보에 관한 다른 법령도 대체로 같은 개념규정을 쓰고 있다.

문제는 '홍길동, 740510-1000000, 서울 강남구 세곡동 100-100'과 같이 그 자체 식별성이 있는 데이터는 아니라 하더라도 다른 결합정보를 충분히 더하고 높은 수준의 데이터처리기법을 활용하면 식별'가능성'은 인정할 수 있는 경우가 많다는 점이다. 통계데이터나 극단적으로 소략한 데이터가 아닌 한 아주 조금이라도 식별가능성이 있게 마련이다. 어떤 경우 개인정보가 되고 어떤 경우 비개인정보인지 기준이 필요해지는 까닭이다. 대법원 판례는 아직 없고, 하급심 재판례로는 이른바 휴대전화번호 뒷자리 사건 판결,¹¹ IMEI 사건 판결¹² 및 최근의 약학정보원 사건 판결¹³이 있다. 휴대전화번호 뒷자리 사건 판결의 경우 휴대전화번호 뒤 네 자리만 있어도 지인(知人)은 누구인지 알 수 있다는 점에서 개인정보라고 하였고, IMEI 사건 판결의 경우 스마트폰의 기기고유번호인 IMEI나 스마트폰 개통과정에서 통신사에 제공되는 USIM 일련번호는 휴대전화 가입신청서 등 가입자 정보와 결합하면 특정인을 식별할 수 있어 개인정보라고 하였다. 특히 IMEI 사건 판결은 결합정보를 획득하기가 매우 어려우라는 점을 거의 고려하지 아니하여, 이른바 절대설에 가깝게, 개인정보를 매우 넓게 파악하는 입장으로 이해된다. 그러나 근래의 좀 더 주류적인 입장은 이른바 상대설을 취하여 취득이 용이한 결합정보만 고려하고 현실적으로 가능성이 높은 분석기술만 고려한다. 약국의 처방전정보를 익명화하여 취득한 것이 문제된 위 약학정보원 사건 판결도 기본적으로 이러한 입장이다.

양자의 구별과 관련하여 이른바 상대설을 취한다 하더라도 추가로 두 가지 점이 문제된다. 첫째, 결합정보와 분석기법을 고려한 식별가능성은 누구를 기준으로 하여야 하는지에 대하여 논란이 있다. 절대설 내지 그에 가까운 견해는 임의의 제3자를 모두 고려하여야 한다는 입장인 반면, 개인정보의 활용을 위하여 오직 개인정보처리자의 관점에서 파악하여야 한다는 입장도 유력하게 주장되고 있다.¹⁴ 그러나 근래에는 개인정보처리자의 관점 이외에 잠재적

08 이상용·최경진, "인공지능과 정보보호법", 인공지능과 법, 2019, 178면.

09 헌법재판소 2005. 5. 26. 선고 99헌마513, 2004헌마190 결정; 2005. 7. 21. 선고 2003헌마282, 425 결정; 2005. 11. 24. 선고 2005헌마112 결정; 대법원 1998. 7. 24. 선고 96다42789 판결; 2011. 9. 2. 선고 2008다42430 판결; 2015. 10. 15. 선고 2014다77970 판결; 2016. 8. 17. 선고 2014다235080 판결.

10 인체유래물의 경우 인체유래물 자체가 개인정보가 되지 아니하는 까닭이다. Mand, "Biobanken für die Forschung und informationelle Selbstbestimmung", MedR 2005, 565, 566.

11 대전지방법원 논산지원 2013. 8. 9. 선고 2013고단17 판결.

12 서울중앙지방법원 2011. 2. 23. 선고 2010고단343 판결.

13 서울중앙지방법원 2017. 9. 11. 선고 2014가합508066, 538302 판결(제1심); 서울고등법원 2019. 5. 3. 선고 2017나2074963, 2074970 판결(항소심).

14 김진환, "개인정보 보호법의 해석 원칙을 위한 제언과 시론", 서울대 법학평론 제3권(2012), 24면 이하; 박혁수, "빅데이터 시대에 개인정보 개념의 재검토", Law & Technology 제10권 제1호(2014), 12면; 이인호, "개인정보 보호법상 '개인정보' 개념에 대한 해석론 -익명화된 처방전 정보를 중심으로-", 정보법학 제19권 제1호(2015), 80면 이하; 정상조, "위치기반서비스 규제에 관한 연구", Law & Technology 제12권 제1호(2016), 18면.

공격자(motivated intruder)가 갖고 있는 결합정보와 동기, 역량도 고려하여야 한다는 견해(절충설)가 유력하다.¹⁵ 위 약학정보원 사건 제1심 및 항소심도 추상적으로는 이와 같은 기준을 사용하였다. 둘째, 식별가능성 유무를 판단할 때에는 기술적인 측면, 즉 통계학적으로 얼마나 정보주체의 식별이 어려워졌는지 여부뿐 아니라, 관리적 측면, 가령 공개의 범위와 방법(일반공개, 제한적 공개, 폐쇄적 이용), 이용목적 및 접근통제권한의 제한, 교육, 접속(log) 기록의 유지 등도 함께 고려하여야 한다.¹⁶ 위 약학정보원 사건도 이러한 측면을 고려하고 있다.

물론 이상과 같은 설명에도 불구하고 여전히 경계 내지 기준이 모호하다는 점은 부정할 수 없다. 위 약학정보원 사건의 경우 제1심과 항소심 모두 청구를 기각하였으나 그 이유는 서로 달랐다. 제1심은 적어도 일부 정보의 경우 복호화가 불가능한 일방향 암호화를 수행한 경우, 정보주체에 관한 암호화된 주민등록번호를 기재한 매칭 테이블을 별도로 제공하였음에도 불구하고 통계분석자료를 생성하여 판매하는 것을 목적으로 하는 상대방로서는 재식별 유인이 없고 실제 복호화를 시도한 바도 없음을 들어 개인정보가 아니라고 하였으나 항소심은 매칭 테이블과 결합하는 경우 복호화가 가능하고 매칭 테이블을 개인정보처리자가 갖고 있었으므로 개인정보에 해당한다면서, 손해가 없다는 이유로 청구를 기각하였다. 이 사건은 상고심 계속 중이므로 대법원의 판단을 기다려보아야 하겠으나, 어떤 결론이 나오든 얼마나 또는 어떻게 식별 가능하여야 개인정보인지에 관한 불명확성이 다 해소되지는 아닐 것이다.

3) 비식별화(de-identification)·익명화(anonymization)

기업이 보유하는 데이터 중 상당부분은 그 본래의 취득목적상 개인정보로 취득된다. 전형적으로 고객 데이터는 그 목적상 식별정보를 포함하고 그 식별정보 아래에 관련 데이터가 연결되어 있는 것이 보통이다. 데이터가 고객 단위로 수집되지 아니하고 물품, 서비스 계정 등을 단위로 수집되고 있다 하더라도 마찬가지이다. 거래의 상대방은 사람일 수밖에 없으므로 물품이나 서비스 계정은 간접적으로라도 특정인에게 연결되어 있게 마련이고, 이 경우 그 데이터는 살아 있는 특정 개인의 행태를 추론할 수 있는 데이터, 즉 식별가능성이 있는 개인정보가 된다. 근로자에 대한 데이터가 식별정보를 포함한 개인정보임은 물론이다. 나아가 전형적인 사물데이터인 공장 등 설비 관련 데이터도, 관련 설비의 관리·이용자가 일정한 규칙이나 표 등으로 관리되고 있는 한, 특정 근로자의 행태에 연결시킬 수 있다.

그중 기본적으로 사물데이터로, 별도의 결합정보를 보태지 아니하면 개인정보가 되지 아니하는 정보는 그 결합정보가 잘 관리되고 있는 한,

¹⁵ 이러한 입장으로 가령, 이동진, “개인정보 보호법 제18조 제2항 제4호, 비식별화, 비재산적 손해 -이른바 약학정보원 사건을 계기로 -”, 정보법학 제21권 제3호(2017), 266면 이하.

¹⁶ 고학수 외, 개인정보 비식별화 방법론 - 보건의료정보를 중심으로 -, 2017, 231면 이하.

결합정보를 가지지 아니하는 제3자에 대한 관계에서는 개인정보가 아니다. 즉 개인정보 개념은 주관적으로도 상대성이 있어 어느 처리자에게는 개인정보인데 다른 처리자에게는 개인정보가 아닐 수 있다. 반면 원정보가 특정 개인에게 귀속되어 수집되는 개인정보의 경우 이를 비개인정보로 바꾸는 조치가 필요한데, 그러한 조치를 일반적으로 비식별화(de-identification)·익명화(anonymization)라고 한다.

두 용어의 사용 내지 개념구분과 관련하여서는 약간의 혼란이 있다. 데이터에서 정보주체를 바로, 즉 그 자체 식별할 수 있는 식별자(identifier)를 삭제하거나 대체하여야 함은 분명하다. 그러나 식별자를 제거하는 것만으로는 식별가능성까지 제거할 수 없으므로 추가적인 조치가 필요한데, 그 정도와 관련하여, 특히 개인정보의 요건으로서 식별가능성을 제거하는 수준에 이르는 경우와 그 수준에는 이르지 못하는 경우를 나누어 각각 어떤 용어를 쓸 것인지 다툼이 있다. 비식별화와 익명화 모두를 전자의 의미로 쓰고 후자는 가명화(pseudonimization)라고 하는 예와, 전자는 익명화, 후자는 비식별화라고 하는 예가 있는 것이다. 우리나라에서는 과거 미국의 용례를 따라 전자를 비식별화라고 하는 것이 보통이었으나,¹⁷ 근래 4차산업혁명위원회는 유럽의 용례를 따라 전자를 익명화, 후자를 가명화라고 하였다.¹⁸

비식별화된 익명화된 식별가능성을 제거하면 더는 개인정보가 아니라는 점에 이론(異論)이 없다. 비식별화·익명화가 개인정보를 비개인정보로 바꾸어 자유로운 활용 가능성을 확보할 수 있는 핵심수단으로 떠오른 까닭이 여기에 있다. 하나 문제는 비식별화 내지 익명화 자체가 개인정보의 ‘가공’에 해당하여 비식별화·익명화를 한다는 점에 대한 정보주체의 동의 등을 받아야 가능하고, 그렇지 아니한 이상 가공된 데이터가 비개인정보가 되는 것은 물론 더더라도, 가공 자체는 불법이 되는 것 아닌가 하는 점이다. 실제로 우리나라와 다른 나라 모두에서 꾸준히 그러한 주장이 제기되어왔다. 그러나 결과적으로 우리나라, 일본, 미국, 유럽 등 거의 모든 곳에서 정보주체의 동의를 받지 아니하여도

¹⁷ 가령 미래창조과학부와 한국정보화진흥원이 2014. 5. 1. 발간한 ‘빅데이터 활용을 위한 개인정보 비식별화 사례집’, 행정자치부와 한국정보화진흥원이 2014. 12. 발간한 ‘개인정보 비식별화에 대한 적정성 자율평가 안내서’ 및 2016. 6. 국무조정실 등 관계부처 합동으로 발표한 ‘개인정보 비식별화 조치 가이드라인 - 비식별 조치 기준 및 지원·관리체계 안내 -’ 등. 이른바 약학정보원 사건의 제1심 및 항소심판결도 이를 ‘비식별화’라고 한다.

¹⁸ 이러한 용어변경은 일부 시민단체들이 익명화가 비식별화보다 식별가능성을 더 강력하게 제거하는, 가령 절대적으로 제거하는 것임을 전제로 비식별화에 반대해왔다는 사정과 관계되어 있다. 그러나 이와 같은 우려 내지 인식은 근거가 없다. 익명은 본래 그리스어에서 유래한 표현으로 한자로든 그리스어로든 식별자를 숨긴다는 이상의 의미가 없고, 유럽에서도 익명화가 것처럼 절대적인 개념이 아니라는 점에는 오늘날 의문이 없기 때문이다. 현행법은 이미 여러 곳에서 익명화 내지 익명처리를 가명화의 의미로 쓰고 있다. 가령 개인정보 보호법 제3조 제7항, 생명윤리 및 안전에 관한 법률 제2조 제19호.

비식별화·익명화하는 것은 가능하다는 것이 일반적인 입장이고, 실무의 태도이기도 하다.¹⁹

비식별화·익명화의 성패는 식별가능성 유무에 달려있다. 그런데 개인정보와 비개인정보의 구별 자체에 불확실성이 있으므로 비식별화·익명화의 성패도 일정한 불확실성을 가질 수밖에 없다. 국무조정실 등 관계부처 합동으로 2016. 6. 발간한 「개인정보 비식별 조치 가이드라인 - 비식별 조치 기준 및 지원·관리체계 안내」는 이와 관련하여 식별자(identifier)는 삭제하고 속성자(attribute value)는 가명처리, 총계처리, 데이터 삭제, 데이터 범주화, 데이터 마스킹 등 세부적인 기법을 적용한 다음, 이른바 k-익명성(k-anonymity) 모델을 활용하여 개인정보 보호책임자 책임 하에 외부전문가가 참여하는 적정성 평가단의 평가를 거치도록 하고, 평가 결과가 '적정'인 경우 활용을 허용하고 있다. 다만 이때에도 불특정 다수에게 제공하는 것은 원칙적으로 금지하고, 데이터의 활용이 끝나면 지체 없이 파기하도록 하며, 비식별정보에 대한 접근통제, 접속기록 관리, 보안 프로그램 설치·운영 등을 요구하고 있다. 이는 개인정보와 비개인정보의 구별이 기술적 측면뿐 아니라 관리적 측면도 함께 고려하여 이루어진다는 점을 반영한 것이다. 이러한 비식별화·익명화 과정은 개별·구체적 위험성 평가 판단을 요하는 이 일의 특성을 고려할 때 기본적으로 타당하나, 실무적으로 일정한 수준의 k-익명성을 확보하면 족한 것처럼 인식될 우려가 있었다. 그밖에 적정성 평가에도 불구하고 그 평가가 잘못된 것일 경우 개인정보가 비개인정보로 전환되지 아니할 가능성이 있음은 물론이다.²⁰ 그러나 절차적으로 적절하게 이루어진 적정성 평가 결과에 따르는 한 처리자의 고의·과실이 부정되어 민·형사책임을 면할 가능성이 높다는 점에서 적정성 평가는 여전히 의미가 있다.

비식별화·익명화를 통하여 비개인정보로 전환된 데이터를 제3자에게 제공하는 경우 제3자가 결합정보를 갖고 있거나 취득하여 정보주체를 식별해낼 위험을 고려할 필요가 있다. 본래의 정보주체에 의한 당초 수집 목적 외의 활용의 경우에도 결합정보의 취득이나 분석기술의 발전으로 정보주체를 다시 식별해낼 가능성이 있다. 이를 재식별(reidentification) 위험이라고 한다. 재식별위험이 커져서 식별가능성이 (개인정보에 포함될 정도로) 높아지면 비개인정보는 다시 개인정보가 된다. 「개인정보 비식별 조치 가이드라인 - 비식별 조치 기준

¹⁹ 이동진(주 15), 262-263면.

²⁰ 이 점에서 위 가이드라인이 비식별조치를 거친 정보는 개인정보로 '추정'한다고 한 것은 엄밀히는 부정확하다. 미국의 경우 HIPAA Privacy Rule은 의료기관 등이 보유하는 건강정보의 비식별화방법으로 Safe Harbor 방식과 전문가 평가 방식을 인정하는데, 어느 것이나 정해진 절차를 거치면 더는 개인정보가 아닌 것으로 간주한다는 점에서 우리 법제와 다르다. 반면 일본과 유럽의 법제는 기본적으로 우리 법제와 같아서, 개인정보가 적절하게 비식별화·익명화되었는지 여부는 언제나 이를 활용 또는 제공하는 개인정보처리자의 책임이고, 어떠한 조치 내지 절차를 거쳐도 그 사실만으로 완전한 보장이 주어지지 않는다는 점이다.

및 지원·관리체계 안내」는 이러한 요인이 생기는 경우 재식별위험을 다시 평가하여야 하고, 제3자 제공을 할 때에는 약정 등으로 재식별 금지와 재제공 또는 재위탁을 제한하며 재식별 위험이 생긴 때에는 제공자에게 통지할 의무를 명시하도록 하고 있다.

어떠한 결합정보와 분석기법을 고려하더라도 재식별위험을 배제할 수 있는 경우가 아닌 한 비식별화·익명화에도 불구하고 일정한 기술적·관리적 조치는 필요하고, 제3자 제공의 경우 당초 정보를 수집하여 조치 후 제공한 개인정보처리자로서는 제공받은 제3자의 손에서 그러한 위험이 실현될 가능성도 고려하여야 하는 것이다.

나. 데이터에 대한 권리

1) 비개인정보인 데이터

데이터거래를 관념하기 위해서는 거래대상인 데이터 이외에 그에 대하여 일정한 권리 내지 지위를 갖는 주체를 상정하여야 한다. 개인정보인 데이터의 경우 개인정보 보호법으로 인하여 추가적인 문제가 생기므로, 그러한 문제가 없는 비개인정보인 데이터부터 본다.

비개인정보인 데이터는 현행법상, 그리고 아직까지는 다른 어느 나라에서도, 일반적 배타적 재산권의 대상이 되지 아니하고 있다. 데이터는 물권의 대상인 물건(민법 제98조)에 해당하지 아니하므로 물권의 일종인 소유권의 대상이 되지 아니한다.²¹ 데이터는 코드이기는 하나 그 내용에 어떤 창작성이 있다고 하기 어렵고 오히려 기왕에 존재하는 정보를 코딩한 것이며, 그 코딩방식에도 어떤 창작성이 없고 기능성·기술적 특성만이 존재하기 때문이다. 특히 데이터거래를 위한 기술적·제도적 기반 중 핵심인 상호운용성(interoperability)의 확보를 위해서는 코딩방식 자체가 어느 정도 표준화되어야 하는데, 그러한 경우 그 표준에 따른 코딩이 독창적 표현이 될 수 없음은 당연하다. 결과적으로 저작권에 의한 보호도 거의 좌절될 수밖에 없다. 데이터의 보호는 아이디어, 즉 내용의 보호가 아닌 표현, 즉 코드 층의 보호이므로 특허권이 문제될 수 없음은 물론이다. 배타적 재산권으로 남는 것은 데이터베이스제작자의 권리 정도인데, 기본적으로 저작권법 및 부정경쟁방지법적 보호를 의도하고 있어 한편으로는 소재를 체계적으로 배열 또는 구성한 편집물일 것을 요구하고 다른 한편으로는

²¹ 상세는 이동진, “데이터 소유권(Data Ownership), 개념과 그 비판”, 정보법학 제22권 제3호(2018) = 고훈수·임용 편 데이터오너십 - 내 정보는 누구의 것인가 -, 2019, 123면 이하. 조심스럽게 이를 긍정할 필요성과 해석상 가능성을 언급하는 것으로 최경진, “데이터와 사법상의 권리, 그리고 데이터 소유권(Data Ownership)”, 고훈수·임용 편 데이터오너십 - 내 정보는 누구의 것인가 -, 2019, 99면 이하, 이를 입법론적으로 인정할 필요가 있다는 것으로 이상용, “데이터 거래의 법적 기초”, 법조 통권728호(2018), 18면 이하. 그러나 이들도 현행법 해석상 인정가능성에 대하여는, 결과적으로, 매우 회의적이다.

데이터베이스의 상당부분을 복제할 것을 요구하므로 이를 비개인정보인 데이터에 적용하기 곤란하다. 오늘날의 데이터 분석기법은 데이터베이스 자체의 체계적 배열이나 구성, 편집을 요구하지 아니하고 데이터의 집적만 있으면 나머지 문제는 분석기법으로 해결하므로, 데이터의 집합 자체가 데이터베이스의 요건을 충족하기 쉽지 아니하고, 데이터베이스의 상당부분의 복제와 데이터의 복제는 구별되는 개념이기 때문이다.²² 일본과 유럽에서의 평가도 이 점에 일치한다.²³

결국 남는 것은 영업비밀 기타 부정경쟁방지법적 보호뿐이다. 이들은, 약간의 논란이 있기는 하나, 기본적으로 사실상의 이익에 대한 불법행위법적 보호이지 법적인 배타적 재산권에 대한 보호는 아니다. 여기에 일반 불법행위법의 보호와 (관련 규정이 적용되는 범위에서) 형법적 보호를 덧붙여도 같다.²⁴ 다만, 영업비밀에 의한 보호는 데이터가 비밀로 유지되어야 가능한데, 일대일 거래의 경우 제공받는 측에 비밀유지의무를 부과하여 영업비밀성을 유지할 수는 있지만 일대다 거래의 경우에는 도저히 비밀로 유지되고 있다고 하기 어려울 수 있다. 일본의 경우 2018. 5. 30. 부정경쟁방지법을 개정하여 이른바 한정제공데이터의 보호에 관한 규정을 도입하였는데, 그 기본취지는 영업비밀로 유지되지 않는 아니하였다 하더라도 (주로 데이터거래 플랫폼을 통하여) 제한적으로 제공되어온 데이터의 무권한 침해에 대하여 영업비밀에 준하는 보호를 제공하겠다는 것이다.²⁵ 그러나 이러한 규정이 없다 하더라도, 데이터 침해는 기계적 복제(dead copy)의 방식으로 이루어지게 마련인데, 그러한 행위는 그 자체 이미 불법행위로 평가할 여지가 있다는 점에서 기본적으로 보호가 전혀 존재하지 아니한다고 하기는 어렵다. 물론, 배타적 재산권을 인정하는 경우처럼 부당하게 취득한 이익을 환수하거나 제3자 추급을 허용하는 효과는 없다는 점에서 불법행위법적 보호에는 한계가 있다.²⁶

데이터에 대한 배타적 재산권이 존재하지 아니하여 결국 불법행위법적 보호의 대상이 될 뿐이므로 국제사법상 불법행위의 준거법, 즉 행위지 및 결과발생지법이 적용된다. 이는 데이터를 사실상 보유·지배하는 자의

²² 우선, 박준석, “빅데이터 등 새로운 데이터에 대한 지적재산권법 차원의 보호가능성”, 고학수·임용 편 데이터오너십 - 내 정보는 누구의 것인가 -, 2019, 159면 이하.

²³ 대표적으로 AI·データの利用に する契約ガイドライン(주 1), 14-15頁; 福岡眞之介·松村英寿(주 4), 75頁 以下; Osborne Clarke LLP, Legal study on Ownership and Access to Data, 2016; Drexl, Designing Competitive Markets for Industrial Data - Between Propertisation and Access, pp. 19 ff.

²⁴ 박준석(주 22), 164면 이하; 이동진(주 21), 128면 이하. 앞의 주 23의 문헌도 대체로 이러한 접근을 지지한다.

²⁵ 우선 박준석(주 22), 168면 이하; 福岡眞之介·松村英寿(주 4), 85頁 以下

²⁶ 이동진(주 21), 134-135면.

영업소재지법이 일응의 기준이 됨을 뜻한다.²⁷

2) 개인정보인 데이터

개인정보인 데이터에 대하여는 개인정보 보호법 등에 의하여 그 정보주체에게 일정한 권리가 부여된다. 이 권리는 기본적으로 인격권의 발현으로 이해되고 있다.²⁸ 그리하여 개인정보의 수집·이용·제공은 모두 정보주체의 동의를 받거나 법률상 근거가 있거나 계약체결 또는 그 이행을 위하여 불가피하게 필요한 경우 등에 한하여 가능하고(제15조 내지 제18조, 뒤에서 다시 본다), 보유기간이 경과하거나 처리 목적을 달성한 경우에는 파기하여야 하며(제21조), 민감정보나 고유식별정보, 주민등록번호의 경우에는 동의 또는 법령상 명시적 근거가 없으면 처리가 불가능하다(제23조, 제24조, 제24조의2)는 등의 제한을 받는다. 또한 개인정보처리자는 개인정보의 분실·도난·유출·위조·변조·훼손을 막기 위하여 이른바 기술적·관리적 조치(Technical and Organizational Measure; TOM)를 다하여야 한다(제29조 이하). 정보주체는 개인정보처리자에 대하여 ① 처리 중인 자신의 개인정보의 (내용)열람을 청구할 권리(제35조), ② 처리하는 개인정보의 정정·삭제를 청구할 권리(제36조), ③ 개인정보의 처리 정지를 요구할 권리(제37조)를 갖고, ④ 같은 법 위반으로 손해를 입은 경우 개인정보처리자에게 손해배상청구를 할 수 있다(제39조, 제39조의2, 민법 제750조 이하). 명문의 규정은 없지만 개인정보자기결정권에 터 잡아 ⑤ 기타 필요한 조치를 청구할 수 있다(일반적 금지청구권, 제51조 참조). 결과적으로 개인정보인 데이터의 경우 그 정보주체에게 광범위한 통제권이 있고, 앞서 비개인정보인 데이터에 대하여 본 데이터의 사실상 지배자의 지위는 널리 정보주체의 통제권에 종속된다. 논란의 소지가 있기는 하나 통설은 개인정보자기결정권이 재산적 측면까지 보호한다고 보지 아니하므로 부당이득반환청구는 허용되지 아니할 가능성이 높다.²⁹

문제는 개인정보인 데이터를 비개인정보로 바꾸거나 개인정보인 데이터를 처리하여 비개인정보인 파생데이터(derivative data)를 생성한 경우 그러한 데이터에 대하여 정보주체가 일정한 통제권을 가지는가 하는 점이다. 전자에

²⁷ 그 이외에 부정경쟁행위의 경우, 경쟁자의 이익만 보호하는 것이 아니라 일반이익으로서 경쟁 자체도 보호하므로 관련시장에 대한 준거법이 고려될 여지가 있다. 우선, Kropholler, Internationales Privatrecht, 5. Aufl., 2004, S. 532 f. 그러나 우리나라에서는 이러한 가능성에 대하여 아직까지 별다른 논의가 없는 상황이다.

²⁸ 정상조·권영준, “개인정보의 보호와 민사적 구제수단”, 법조 통권630호(2009), 15면 이하.

²⁹ 이 점에서는 하나의 정보가 복수의 개인과 관련을 맺을 수 있어 특정인에게 배타적으로 할당될 수 없는 경우가 적지 아니하다는 점을 고려할 필요가 있다.

대하여는 정보주체의 통제권을 인정하지 아니하는 것이 일반적인 견해이지만,³⁰ 후자의 경우에는 파생데이터 생성에 대하여 정보주체의 동의를 받아야 한다.³¹ 각종 인체유래물연구 등이 정보주체의 동의하에 이루어지는 이유이기도 하다. 동의를 받아 이미 비개인정보인 데이터를 생성하였는데 이후 동의가 철회된 경우 그 철회의 효력이 비개인정보인 파생데이터까지 미치지 아니한다고 보는 것이 일반적인 이해이자 실무이나, 역시 논란의 소지가 있다.

개인정보자기결정권이 인격권에 해당하는 한 그에 대한 준거법도 불법행위의 준거법규정에 따른다.³²

3. 데이터거래의 구조와 쟁점

가. 데이터거래의 구분

1) 데이터제공·데이터창출계약과 데이터플랫폼규약

주로 B2C거래이고 따라서 소비자보호가 문제되는 디지털콘텐츠거래와 달리 B2B거래가 주인 데이터거래의 구체적 형성과 쟁점에 관한 연구는 아직까지 활발하다고 할 수 없다. 이 점에서 일본 경제산업성이 각국의 데이터거래에 관한 연구와 자료를 분석하여 2018. 6. 발표한 「AI-데이터의 이용에 관한 계약 가이드라인 데이터편」(AI-データの利用に する契約ガイドラインデータ編)이 참고가 된다. 위 가이드라인은 데이터이용계약을 데이터제공형계약, 데이터창출형계약 및 데이터공용형계약의 세 유형으로 나누고, 각각의 쟁점을 설명한 다음, 데이터제공형계약과 데이터창출형계약의 모델계약서안을 제시하고 있다.³³ 그 이외에 정부가 제시한 것은 아니나 각종 데이터거래계약 사안(私案)이

³⁰ 그러나 적어도 행태데이터는 비개인정보여도 배타적 재산권의 대상이 되어야 한다는 견해로, Fezer, Repräsentatives Dateneigentum. Ein zivilgesellschaftliches Bürgerrecht, 2018, 특히 S. 35 ff.. 이는 물론 입법론이다.

³¹ 박광배·채성희·김현진, “빅데이터 시대 생성정보의 처리 체계 - 추론된 정보의 처리에 관한 우리 개인정보보호법의 규율과 개선 방안에 관한 고찰 -”, 정보법학 제21권 제2호(2017), 181면 이하. 다만 같은 문헌은 개인정보를 비개인정보로 만든 뒤 그로부터 정보를 생성하는 경우에도 같다는 취지로 보이나, 완전한 비개인정보로 만들 수 없다는, 즉 비식별화가 불가능하다는 취지라면 별론, 비식별화된 뒤의 생성이라면 타당하다고 할 수 없다. 같은 문헌이 어떤 경우를 상정하여 또는 어떤 논리 하에 위와 같은 결론에 이르렀는지는 반드시 분명하지 아니하다.

³² Kropholler(주 27), S. 529 ff. 피해자의 상거소지법 적용을 중심으로 한 인격권 침해 및 명예훼손에 대한 준거법결정에 관한 여러 접근은 석광현, 국제사법해설, 2013, 409- 410면. 우리 국제사법상 어떻게 처리되어야 하는지에 대하여는 직접 언급하지 아니한다.

³³ AI-データの利用に する契約ガイドライン(주 1), 65頁 以下. 그 이외에 COMMISSION STAFF WORKING DOCUMENT on the free flow of data and emerging issues of the European data economy. Accompanying the document Communication. Building a European data economy, 2017 {COM(2017) 9 final}, pp. 12 ff.

발표되고 있고,³⁴ 거래 가이드라인도 있다.

데이터거래에도 계약자유 원칙이 적용된다. 그러나 크게 몇 가지 유형으로 나누어보는 것은 가능하고 의미가 있다. 유럽연합은 오픈 데이터 접근(open data approach), 데이터 시장에서의 금전화(data monetisation on a data marketplace), 폐쇄된 플랫폼에서의 데이터 교환(data exchange in a closed platform)으로 나누고,³⁵ 일본 경제산업성 가이드라인은 데이터제공형, 데이터창출형 및 데이터공용형계약으로 나눈다. 데이터제공형계약은 데이터의 양도, 이용허락 및 공동이용(상호이용허락)을 포괄한다. 이들은 특정 데이터에 대하여 어느 한 사람(제공자)이 상대방에게 데이터를 제공하는 것을 핵심내용으로 한다는 점에서 공통적이다. 가장 흔한 유형인데, 그중에서 데이터제공자가 데이터를 보유하면서 상대방에게 데이터 사용을 허락하고 제공하는 이용허락이 특히 보편적이다. 데이터창출형계약은 필요에 맞추어 장비(센서)를 부착하여 데이터를 생성하거나 기존 데이터를 가공하여 제공하는 것을 말한다. 일본 가이드라인에서는 별다른 주목을 받지 못하였으나, 복수의 데이터를 결합하여 새로운 데이터를 제공하는 경우도 이에 해당할 것이다. 마지막으로 데이터공용형계약은 데이터거래·공유 플랫폼에서 다수의 제공자와 수요자가 규약에 따라 거래하는 것을 말한다. 유럽연합의 오픈 데이터 접근과 데이터 시장에서의 금전화는 이중 데이터제공형 및 데이터창출형을 포괄한다.

이 글에서는 위 두 구별방식 중 각각 마지막 플랫폼을 제외하고, 앞의 두 유형만 다룬다. 마지막 유형은 다자간 플랫폼 내 거래라는 점에서는 성질을 달리하나 그 외에는 앞의 두 유형의 연장선상에 있고, 오히려 플랫폼 디자인이 문제되어 접근 관점에 차이가 있기 때문이다. 하나 지적할 특징 내지 고려요소는, 특히 공유 플랫폼에 업로드 되거나 공유 플랫폼에서 제공되는 데이터는 영업비밀로 보호되지 아닐 가능성이 있다는 사실이다. 일본의 개정 부정경쟁방지법이 이른바 한정제공데이터의 보호에 관한 별도의 규정을 마련한 까닭이 여기에 있다. 바꾸어 말하면 한정제공데이터의 보호에 관한 일본

³⁴ 가령 長谷川俊明 編 データ取引契約の基本と書式, 2018, 80頁 以下; 福岡眞之介·松村英寿(주 4), 320頁 以下. 표준조항을 제공하는 것은 아니나 그러한 접근이 바람직하다는 지적은 매우 많다. 가령 Heckmann, Big Data im Freistaat Bayern. Chancen und Herausforderungen, 2015, S. 132. 유럽연합 차원에서도 표준계약서의 활용 가능성이 제안되고 있다. COMMISSION STAFF WORKING DOCUMENT(주 33), pp. 30 f.

³⁵ COMMISSION STAFF WORKING DOCUMENT(주 33), pp. 5 ff.

부정경쟁방지법 개정은 주로 공유플랫폼을 염두에 둔 것이다.³⁶

2) 비개인정보인 데이터의 거래와 개인정보인 데이터의 거래

다른 한편 비개인정보인 데이터의 거래와 개인정보인 또는 개인정보였던 데이터의 거래도 구별할 필요가 있다. 앞서 본 바와 같이 개인정보에 대하여는 여러 규제가 있고, 이는 데이터거래에도 상당한 영향을 미친다. 또한 개인정보가 비식별화·익명화되어 비개인정보가 되었다 하더라도 재식별위험이 존재하는 한 다시 개인정보 보호법의 규제를 받을 가능성이 있다. 이는 데이터거래 과정에서도 배려되어야 한다. 거래의 대상인 데이터가 어떤 성질을 갖는지에 따른 구별이 불가피한 까닭이다.

나. 데이터거래 일반

1) 데이터거래의 일반적 고려요소

가) 데이터의 처분가능성

데이터거래에서 먼저 문제되는 것은 문제가 된 데이터가 거래가능한가 하는 점이다. 어떤 하나의 재화를 보유하고 이용할 수 있는 권리자라는 사실이 곧바로 그 데이터를 양도하거나 가공하거나 제3자의 이용을 허락할 수 있음을 뜻하지는 아니하고, 이는 배타적 재산권이 설정되어 있지 아니한 데이터의 경우에도 마찬가지이다.

일반적으로 데이터를 제공하고자 하는 측이 직접 그 데이터를 수집하거나 생성하였다면 그에 대하여 사실상의 처분권을 가진다고 볼 수 있다.³⁷ 그러나 데이터를 제공하고자 하는 측(재제공자)도 그 데이터를 제3자(원제공자)로부터 제공받은 것일 때에는 약정에 의하여 전(轉)매도 또는 전대(轉貸)가 금지되어 있을 가능성이 높다. 위와 같은 약정은 일응 채권적 효력만을 가지므로 전(轉)매도 또는 전대(轉貸)하더라도 원제공자에 대한 재제공자(원수령자)의 계약위반이 될 뿐 재수령자에게는 별 효력이 없다. 데이터에 대한 배타적 재산권이 존재하지 아니하는 한 원제공자가 재수령자에게 직접 데이터의 '반환' 또는 삭제를 요구하기도 어렵다. 다만 당해 데이터가 '공공연히 알려져 있지 아니하고 독립된

³⁶ 구체적인 문례로는 福岡眞之介·松村英(주 4), 417頁 以下. 플랫폼 멤버십에 관한 규정, 멤버십이 인정되는 한 당연히 이용할 수 있는 공유데이터에 관한 규약, 운영자의 권한과 의무 및 책임 또는 그 배제에 관한 조항 등이 주를 이룬다. 그밖에 플랫폼 멤버에 대하여는 직접 거래 상대방이 아니어도 규약에 따른 권리관철을 인정함으로써 데이터에 대한 배타적 재산권에 준하는 권리를 창설할 가능성이 생긴다. 특히 데이터공유플랫폼의 역할과 기능에 대하여는 Richter and Slowinski, "The Data Sharing Economy: On the Emergence of New Intermediaries", ICC 50 (2019), 4 ff.

³⁷ 이동진(주 21), 131-133면; Specht, Konsequenzen der Ökonomisierung informationeller Selbstbestimmung: Die zivilrechtlicher Erfassung des Datenhandels, 2012, S. 139 f.

경제적 가치를 가지는 것으로서, 비밀로 관리된 정보'에 해당하여 영업비밀로 보호되고 있었던 경우, 이를 알고 데이터를 취득한 재수령자도 영업비밀을 침해한 것이 되고, 데이터 취득 후 알게 되어도 그 사용이 금지된다(부정경쟁방지 및 영업비밀보호에 관한 법률 제2조 제2호, 제3호 라.마.목).³⁸ 즉 이 경우에는 원제공자가 재수령자를 상대로 폐기 등을 요구할 수 있고,³⁹ 심지어는 형사처벌로 제재될 수도 있는 것이다. 그밖에 데이터의 제공이 다른 재산권과 관계하는 경우, 가령 데이터의 표현 또는 내용이 저작권, 특허권 등으로 보호받고 있거나 데이터제공 매체의 소유권이 문제되고, 데이터제공자가 저작권자, 특허권자 또는 매체소유자가 아닌 경우 데이터제공이 필연적으로 저작권·특허권·소유권침해를 수반하게 되는지 따져보아야 한다. 소유권은 별론, 그 이외의 권리는 데이터를 제공함으로써 침해될 가능성이 생기므로, 제공하는 경우에도 약정에 의하여 이용목적과 방법을 제한함으로써 위 권리가 침해될 가능성을 피하여야 한다. 그러한 조치를 취하지 아니한 경우는 물론, 그러한 조치를 취한 경우에도 데이터수령자가 약정을 위반하여 저작권·특허권 등을 침해하였다면 데이터제공자는 그에 대한 지식재산권 등 침해 방조책임(민법 제760조 제2항)을 질 가능성이 있다.

나) 데이터거래와 경쟁·경쟁법

거래 가능한 데이터라 하더라도 그 데이터가 거래에 적합한 것인가 하는 점은 별도로 검토할 필요가 있다. 데이터제공에 양도와 (일방적 또는 상호)이용허락이 있다는 점은 이미 본 바와 같다. 그중 양도는 데이터제공자는 더는 그 데이터를 쓰지 아니하는 것을 예정하므로 전혀 다른 결정에 터 잡고 있지만, 이용허락은 데이터제공자와 데이터수령자가 모두 데이터를 이용할 것을 전제하므로 양자 사이의 경쟁에 일정한 영향을 미친다. 데이터창출은 존재하지 아니하는 데이터를 만들어내는 것을 포함하지만 그 또한 경쟁에 일정한 영향을 미치게 마련이다. 데이터거래가 필요한 까닭에는 데이터가 품질손상 없이 거의 무제한으로 복제할 수 있고 복제에 비용도 (거의) 들지 아니한다는 점 이외에, 그럼에도 불구하고 큰 가치를 창출할 수 있을 만큼 충분히 많은 양의 데이터를 수집하는 것은 매우 어렵다는 점이 포함된다. 가령 공장의 설비운용 데이터, 차량 운행 데이터, 특정 시설의 이용 데이터 등 상당수의 리얼 데이터(real data)가 정부·지방자치단체, 소수의 기업에 의하여 운용되고 있어, 데이터를 수집할 수 있는 입구(gate)에 사실상 그들만이 접근할 수 있기 때문이다. 개인정보의 경우에도 마찬가지이다.

³⁸ 판례는 영업비밀의 부정취득에 대하여 경쟁관련성 및 실제 사용 여부를 따지지 아니한다. 대법원 2011. 7. 14. 선고 2009다12528 판결.

³⁹ 가령 조성필, "부정경쟁방지 및 영업비밀보호에 관한 법률상의 영업비밀의 보호에 관한 연구 - 민사적 구제 및 이에 대한 판례를 중심으로 -", 광주지방법원 재판실무연구 2000(2001), 301면 이하.

개개의 병·의원 이용 데이터는 개개의 병·의원에, 건강보험요양급여청구 데이터는 국민건강보험공단과 건강보험심사평가원에, 금융거래내역은 비교적 소수의 금융기관들에, 그리고 오늘날 가장 많은 이용자 행태정보를 수집하고 있는 인터넷환경에서는 소수의 플랫폼사업자들에게 각각 집적된다. 그리고 이들에게는 바로 이러한 정보의 집적 자체가 경쟁력의 주요 원천 중 하나이다. 개개의 데이터제공자로서는 당해 데이터를 데이터수령자에게 제공하는 것이 자신의 경쟁상의 우위를 위협하지 아니하는지를 검토하지 않을 수 없는 것이다.⁴⁰

그러나 경쟁상 우위의 고려에 법적 한계가 있을 수 있다. 오늘날 데이터경제의 대두와 관련하여 논의되는 한계로는 특히 다음 둘이 두드러진다. 첫째, 데이터제공자가 당해 시장에 진입하여 경쟁하는 데 필수적인 데이터를 사실상 독점하고 있어 데이터제공이 없으면 시장진입 및 경쟁이 봉쇄되는 경우 이른바 필수설비이론(essential facility doctrine)에 준하여 경쟁법상 데이터이용허락을 강제할 수 있다는 주장이 꾸준히 제기되고 있다.⁴¹ 둘째, 데이터제공자가 자신의 데이터수령자에게 긴요한 데이터의 제공을 조건으로 데이터를 제공하기로 하거나 구속조건부 거래를 하는 경우 그것이 거래상지위남용 기타 불공정거래행위(독점규제 및 공정거래에 관한 법률 제23조)가 될 수 있다.⁴² 이들은 어느 것이나 아직 형성 중인 법리로, 어떠한 경우에 법 위반이 인정될지는 분명하지 아니하다. 그러나 사실상 데이터를 독점하는 기업이 데이터제공에 매우 소극적이거나 데이터거래를 기화로 부당한 거래조건을 관철시킨다면 위법 여부와 관계없이 적어도 사회적 비난을 받을 소지는 상당하다는 점에 비추어도 고려요소 중 하나임은 부정하기 어렵다.

다) 데이터의 경제적 가치

데이터를 거래하는 경우에는, 일부 무상거래를 제외하면, 대체로 유상거래가 되기 쉽다. 이때에는 적절한 대가를 정하는 문제가 생긴다. 그런데 데이터는 일종의 원료에 해당하여 그로부터 어느 정도의 가치를 창출할지는 각각의 이용목적과 역량에 의존하는 바가 크고, 개개의 데이터의 개성 내지 편차도 상당하며, 그 자체 비용 없이 복제가 가능하고, 특히 데이터거래시장도

40 福岡眞之介·松村英寿(주 4), 29頁 참조.

41 홍대식, "데이터 소유권(Ownership)을 둘러싼 법적 쟁점과 경쟁법 원칙의 적용", 고학수·임용 편 데이터오너십 - 내 정보는 누구의 것인가 -, 2019, 192면 이하. 이 문제에 대한 독일과 프랑스 경쟁당국의 공동 policy paper로, Autorité de la concurrence and Bundeskartellamt, Competition Law and Data, 2016. 또한 Drexl(주 23), pp. 41 ff.; Dewenter/Lüth, Datenhandel und Plattformen, 2018, S. 56 ff. 일본의 논의는 福岡眞之介·松村英寿(주 4), 303-304頁.

42 福岡眞之介·松村英寿(주 4), 301-302, 304頁 以下.

활성화되어 있지 아니하여 '객관적' 가치산정이 매우 까다롭다.⁴³ 데이터 가치산정의 노후를 발전시킬 필요가 있음은 물론이나, 가치산정이 특히 곤란한 경우 대안으로 데이터를 이용하여 데이터수령자가 얻은 이익에 지분으로 참여하는 모델, 데이터와 데이터를 교환하는 모델(상호이용허락)을 생각해볼 수 있다.

2) 데이터거래계약의 주요 규율대상

가) 적용되는 법

데이터거래계약에는 어떠한 법이 적용되는가. 데이터는 가상의 재화로 그 물리적 소재지를 이야기하기 어려운 측면이 있으므로, 계약에 대한 준거법을 지정할 필요가 있다. 국제사법상 계약의 준거법은 당사자 자치에 맡겨져 있으므로, 약정으로 준거법을 정하면 된다. 이하 이 글에서는 일응 한국법이 준거법으로 지정되거나 준거법인 경우를 전제하기로 한다.

데이터거래계약을 다루는 전형계약이나 법률규정은 존재하지 아니한다. 다른 나라에서는 그와 같은 입법의 필요성도 논의되고 있으나⁴⁴ 아직까지 별 호응을 얻지 못하고 있다. 무엇보다도 계약자유의 원칙과 기존의 계약유형으로 대부분 대응할 수 있다는 전망 때문이다.⁴⁵

데이터거래계약은 어떻게 성질결정(qualification)될 수 있는가. 이 점은 임의규정의 적용과 관련하여 중요한 의미를 갖는다.

먼저 분명히 할 점은 데이터는 배타적 재산권의 대상이 아니고, 따라서 데이터에 대한 배타적 재산권의 이전 또는 이용에 대한 동의(Einwilligung)의 의미에서 그 이용허락도 관념하기 어렵다는 것이다. 그럼에도 불구하고 데이터 '양도'라는 표현을 쓰는 것은 데이터에 대한 사실상 지배의 이전이 가능하기 때문이다. 데이터'양도인'과 '양수인' 사이에서 데이터'양도인'이 더는 데이터의 이용 및 그 처분에 대하여 종국적 권한을 가지지 아니하고 데이터'양수인'이 데이터의 이용 및 처분에 대하여 종국적 권한을 가지기로 하면서, 데이터에 대한 사실상 지배를 데이터'양수인'에게 넘기는 경우 두 사람 사이에서는 배타적 재산권 양도와 같은 효과가 생긴다. 이는 데이터가 저장되어 있는 (유일한) 매체를 (물권법적으로) 양도하는 방법 이외에 데이터를 '양수인'에게 전송(傳送)한 다음 '양도인' 측에 남은 데이터를 삭제하는 방법으로도 가능하다. 인도(引渡, 민법

43 가령 이상용(주 21), 12면. 경제학적 관점에서 데이터의 가치산정에 대한 개관으로는 Jentzsch, "Datenhandel und Datenmonetisierung: Ein Überblick", Stiftung Datenschutz (hrsg) Dateneigentum und Datenhandel, S. 177 ff. 같은 문헌도 지적하고 있는 것처럼 미국에서 데이터거래가 활발하게 이루어지는 데는 데이터브로커, 데이터릴러에 의한 시장조성과 가치평가의 기여가 적지 아니하다.

44 가령 COMMISSION STAFF WORKING DOCUMENT(주 33), pp. 31 ff.

45 이 점은 디지털콘텐츠계약에 대하여도 그러하다. Grünberger. "Verträge über digitale Güter", AcP 218 (2018), 213, 214 f.에 소개된 각국의 논의현황 참조.

제196조 제1항)는 전점유자의 사실상 지배를 배제하고 현점유자의 사실상 지배를 확보함으로써 달성될 수 있고, 배타적 재산권이 존재하지 아니하는 한 인도만으로도 재산적 이익이 넘어갔다고 할 수 있기 때문이다.⁴⁶ 이처럼 데이터를 넘겨주는 행위는 사실행위이고 법률행위·처분행위는 아니지만, 채권법적으로 위 계약을 유상인 경우 (금전과 교환하면) 매매나 (데이터 등과 교환하면) 교환, 무상인 경우 증여로 파악하는 데 큰 장애는 없다.⁴⁷ 특히 우리 민법은 유상계약 일반에 대하여 매매에 관한 규정을 준용하고 있다(민법 제567조). 데이터양도계약에 대하여 매매법을 적용함으로써 매매에 관한 임의규정, 특히 담보책임법을 적용할 수 있게 된다. 데이터를 '양수'하였는데, 이후 원제공자가 부정경쟁방지 및 영업비밀보호 등에 관한 법률상 금지청구권을 행사하여 데이터를 쓸 수 없게 되면 타인권리의 매매에 준하여 담보책임을 물을 수 있고(민법 제570조 이하), 제공받은 데이터에 흠이 있으면 하자담보책임을 물을 수 있다(민법 제580조).⁴⁸

그러나 좀 더 흔한 유형은 데이터이용허락이다. 이는 데이터제공자가 여전히 데이터를 제공하면서 데이터수령자도 이용할 수 있게 해주는 것으로, (비기술적·비법적 의미의) 데이터공유에 해당한다. 다른 나라에서도 data sharing agreement, Datennutzungsvertrag, contrat de licence de données 등의 이름으로 이 유형의 계약이 중점적으로 검토, 이용되고 있다. 데이터이용허락은 그 표현에서도 알 수 있듯 데이터공유가 지식재산권 이용허락(licence)과 비슷하다는 점에 착안한 것이다.⁴⁹ 데이터는 배제성·경합성이 없어 복제와 동시이용이 가능하므로 이용을 허락하면 된다는 생각이다. 다만, 두 가지 점에 주의하여야 한다. 지식재산권에 대하여는 배타적 재산권이 있으므로 채권적 이용허락과 물권적·배타적 이용허락을 구분할 수 있으나, 데이터는 현행법상 배타적 재산권의 대상이 아니므로 채권적 이용허락을 생각할 수 있을 뿐이다. 채권적 이용허락에서 비독점적 이용허락과 독점적(배타적) 이용허락을 구분하는 것은 가능하지만 말이다.⁵⁰ 또한 데이터는 배타적 재산권의 대상이 아니므로 이미 이용 가능한 정보에 대하여 법적으로 이용을 허락(Einwilligung)하는 것이 아니라 데이터를 이용할 수 있는 상태로 수령자에게 제공하되 데이터제공자도 데이터를 사용할 (종국적) 권리를 보류하는

46 나아가 당사자 사이에서 '인도인'의 사실상 지배 또는 이용을 '인수인'이 허락하는 범위로 제한하고 종국적으로는 '인수인'에게만 귀속하도록 하는 특약이 있는 경우에도 양도와 모순되지 아니한다고 볼 수 있다.

47 광윤직 편집대표 민법주해[XVI], 1997, 109-112면(남호순 집필부분); Specht(주 37), S. 125 ff.; Haedicke, Rechtskauf und Rechtsmängelhaftung, 2003, S. 94.

48 Specht(주 37), S. 190 ff.

49 福岡眞之介·松村英寿(주 4), 48-49頁.

50 Specht(주 37), S. 161 ff.

방식으로 이를 재구성하여야 한다. 지식재산권 이용허락과 경제적·사실적 구조는 비슷하나 그 법적 구성에는 차이가 있는 셈이다. 결국 데이터이용허락은 지속적이든 기간제한이 있든, 목적의 제한이 있든 없든, 데이터수령자의 이용이 제한되어 있고, 그와 동시에 또는 제한을 넘는 한 잔여 이용권 및 처분권이 데이터제공자에게 유보된 경우를 가리킨다. 이때 데이터를 제공해주기만 하고 데이터제공자 자신은 데이터를 이용하지 아니하기로 한다면 임대차·사용대차에, 데이터제공자와 수령자가 동시에 데이터를 사용한다면 지식재산권 이용허락 유사의 독자적인 계약유형에 해당한다.

반면 데이터창출은 도급의 성격을 가진다. 데이터수령자의 주문(注文)에 따라 데이터를 생산하여 공급하여야 하기 때문이다. 데이터에 흠이 있는 경우 그에 대한 담보책임도 도급법 규정(민법 제667조)에 의한다.

데이터양도의 경우 양도인이 데이터를 폐기하고, 데이터이용허락의 경우 이용기간종료 후 데이터수령자가 데이터를 폐기하여야 한다. 인도·재인도와 달리 폐기가 제대로 이루어졌는지 확인할 필요가 있다. 이들 데이터거래계약은 폐기의무와 함께 폐기 여부를 확인하는 절차 등에 대하여도 규정할 필요가 있다.

나) 거래대상과 범위의 특정

데이터거래계약도 급여 목적물을 특정하여야 한다. 구체적으로 어떠한 데이터가 거래대상이 되는지 특정할 필요가 있다. 데이터의 내용과 수량 이외에 데이터의 포맷(format)과 제공방법도 특정하여야 한다. 제공방법이 전기통신이고, 데이터가 계속적으로 제공되는 경우 제공방법에 대하여도 구체적으로 합의하는 것이 바람직하다.

나아가 데이터가 영업비밀로 보호되고 있고 계속 보호되어야 하거나, 영업비밀은 아니지만 제3자의 접근을 사실상 제한할 필요가 있는 한,⁵¹ 데이터수령자에게 데이터거래계약에 의하여 비밀유지의무를 부과하여야 한다. 데이터양도나 데이터창출처럼 데이터제공자에게 데이터가 남지 아니하는 경우 및 오픈 데이터는 별론, 그 이외에는 이러한 조치가 요구되는 경우가 대부분일 것이다. 당해 데이터에 대한 사실상 지배권이 종국적으로 데이터제공자에게 보류되기를 원한다면 데이터수령자가 법적으로 요구되는 수준의 비밀유지를 해주어야 하므로, 일반적 비밀유지의무 이외에 접근제한, 접근권자의 제한, 접근(log)기록의 관리 등 구체적인 조치를 넣고 계속적 협의·보완도 가능하게 하는 것이 좋다. 가령 관련 사항에 대하여 데이터제공자가 확인 내지 감독할 권한을 포함시키고, 유출 등 사고 발생 시 지체 없이 데이터제공자에게 통지하는

51 거래대상으로서 희소성을 위해서는 그렇게 관리할 필요가 있을 것이다. 일본의 개정 부정경쟁방지법상 한정제공데이터도 관리를 요구한다.

등 협력의무를 구체적으로 정할 수 있다.⁵²

그밖에 데이터의 표현이나 내용, 저장매체가 저작권·특허권·소유권의 대상인 경우 데이터거래가 그러한 권리에 대하여 미치는 영향도 분명히 하는 것이 바람직하다. 이는 뒤에 보는 이용관계와도 관계된다. 이용목적이 적절히 제한되면 저작권·특허권 등과는 무관해질 수도 있는 것이다. 그러한 경우 저작권·특허권은 여전히 데이터제공자에게 유보되고 데이터제공이 그에 대한 이용허락을 뜻하지 아니한다는 점을 명시함이 바람직할 수 있다.

대가로는 금전 이외에 데이터 제공(교환), 데이터수령자의 이득참여 등 다양한 형태가 고려될 수 있다.⁵³ 단지 데이터 제공을 위한 처리 수수료 정도만 받고 제공 내지 사실상 공개하는 데이터도 생각할 수 있는데, 이때 데이터거래 부분은 무상계약이다. 다만, 이때에도 데이터에 관한 권리가 많은 부분 법적으로 규율되어 있지 아니하여 계약으로 규율할 수밖에 없으므로, 데이터거래계약은 여전히 필요하다.

다) 이용관계 및 그 결과에 대한 규율

데이터양도나 데이터창출의 경우에는 덜 문제되지만, 데이터이용허락의 경우에는 이용목적 제한하는 것이 보통이다. 전형적인 방식은 데이터거래계약이 정한 목적에 한하여 이용할 수 있고, 목적 외 이용이나 제3자제공은 허용하지 아니하며, 나아가 이용관계의 기간을 제한하고 갱신조건을 정하거나, 무기로 약정한 경우 이용관계의 종료, 즉 계약해지사유를 정하는 것이다.⁵⁴ 현재 거래계에서 사용하는 계약이 대체로 재사용(re-use)을 허용하지 아니하는 데 대한 비판이⁵⁵ 있으나, 데이터가 경쟁에 미치는 영향을 고려할 때 (뒤에 보는 개인정보에서와 마찬가지로) 데이터수령자의 이용을 엄격하게 통제하는 것이 불가피한 경우가 많다.

좀 더 중요한 것은 이 데이터를 이용하여 발생한 파생데이터나 파생지식재산권의 권리귀속을 정하는 것이다. 이러한 권리의 귀속은 당사자 사이의 약정으로 정할 수 있는 반면, 정하지 아니하면 법률관계가 모호해지게 마련이기 때문이다. 파생데이터나 파생지식재산권이 데이터제공자에게도 중요하다는 등의 사정이 있다면 데이터제공자에게 이를 귀속시키거나 공유로 정하고 지분에 따라 수익을 배분할 수도 있다. 어느 경우든 데이터수령자는

⁵² 다만, 데이터가 포괄적으로 이전되었는데, 그중 이미 보유하고 있던 데이터가 포함되거나 제3자에게 동일한 데이터를 받거나 상대방에게서 따로 받거나 직접 생성하였거나 공지(公知)의 데이터가 포함된 경우에는 해석상 혼란을 피하기 위하여 비밀유지의무에서 제외하는 규정을 둘 수 있다. 福岡眞之介·松村英寿(주 4), 68-69頁.

⁵³ 福岡眞之介·松村英寿(주 4), 65頁 以下.

⁵⁴ 福岡眞之介·松村英寿(주 4), 57頁 以下.

⁵⁵ COMMISSION STAFF WORKING DOCUMENT(주 33), pp. 16 f.

파생데이터나 파생지식재산권을 필요하여 생성한 것이므로, 적어도 이용권한은 부여할 필요가 있을 것이다.⁵⁶

라) 책임배분

데이터거래계약에 의하여 제공된 데이터가 필요한 수량과 품질에 미치지 못하거나 요구되는 포맷(format)에 맞지 아니하여 아예 읽어 들일 수 없다는 등의 사정이 있으면 하자담보책임 이외에 계약책임(민법 제390조 등)을 진다는 데 이론(異論)이 없다. 그러나 데이터에 요구되는 내용과 수량, 품질을 어떻게 특정할 것인가 하는 점은 제법 까다롭다. 데이터는, 특히 다량인 경우, 완전하지 아니한 것이 보통이므로 어느 정도의 불완전성까지 용인되는 것인지를 따져야 하기 때문이다. 이미 존재하는 데이터의 경우 일부를 확인한 뒤 현존하는 상태대로 거래할 수 있겠지만, 데이터를 가공하거나 창출해야 하는 경우에는 논란의 소지가 있다. 데이터수령자의 이용목적과 요구한 데이터가 서로 부합하지 아니하여 결과적으로 데이터수령자가 원하는 성과를 거두지 못한 경우 데이터제공자의 책임인지, 아니면 데이터수령자의 책임인지도 분명히 함이 바람직하다. 가령 현존하는 데이터의 제공의 경우 데이터의 흠으로 인한 오작동 기타 문제에 대하여는 데이터수령자의 책임으로 한다는 규정을 둘 수 있는 것이다. 데이터창출의 경우에는 그러한 흠이 도급인의 잘못되거나 부정확한 지시에서 비롯하였는지, 수급인의 부적절한 일에서 비롯하였는지에 따라 책임관계가 달라질 것이다. 한편, 데이터수령자가 당해 데이터로 어떠한 행위를 하였는데, 데이터의 흠으로 말미암아 그 행위가 제3자에게 해를 끼친 경우 그 책임을 데이터제공자가 질 것인지, 데이터수령자가 질 것인지 정하여 두는 것이 바람직하다. 물론 그러한 약정은 내부적 구상관계에 영향을 줄 뿐, 대외적 책임에 직접 영향을 주는 것은 아니다.⁵⁷

그밖에 데이터에 대하여 제3자가 부정경쟁방지법상 영업비밀 침해금지청구권을 주장하거나 저작권·특허권 등을 주장하여 데이터수령자가 사실상 데이터를 쓸 수 없게 되는 경우 등에 대하여는 데이터제공자가 담보책임을 지는 것이 원칙이다. 그러나 약정으로 책임을 배제할 수 있음은 물론이다. 반대로 데이터수령자의 약정에 반하는 이용·제3자제공으로 인하여 데이터제공자가 (원)권리자에 대하여 책임을 지게 되는 경우 내부적으로는 데이터수령자가 책임을 인수한다고 약정하는 것이 바람직할 것이다.

⁵⁶ 福岡眞之介·松村英寿(주 4), 70頁 以下.

⁵⁷ 福岡眞之介·松村英寿(주 4), 61頁 以下.

다. 개인정보인 또는 개인정보였던 데이터거래

1) 개인정보의 제3자 제공

가) 개인정보 제3자 제공의 요건과 절차, 방법

개인정보인 데이터를 거래의 대상으로 삼을 때에도 도대체 거래가 가능한가 하는 점을 먼저 검토하여야 한다. 이는 개인정보 보호법상으로는 목적 외 제3자 제공의 문제이다.

개인정보를 수집한 개인정보처리자가 목적 외 제3자 제공을 하기 위해서는 개인정보 보호법 제18조 제2항 각호 중 어느 하나의 사유가 있어야 한다. 그중에서 공공기관이 법령상 제3자 제공을 받는 경우를 제외하면 현실적으로 고려의 대상이 되는 것은 제1호와 제4호 정도이다.

먼저 개인정보 보호법 제18조 제1항 제1호는 정보주체로부터 별도의 동의를 요건으로 목적 외 제3자 제공을 허용한다. 이러한 동의는 수집 및 목적 내 처리에 대한 동의와 구별되어야 하고, 제3자 제공에 대한 동의를 하지 아니한다는 이유로 본래의 계약의 체결을 거부하여서는 안 된다(개인정보 보호법 제22조). 그러나 이는 본래의 계약의 체결을 거부하여서는 안 된다는 뜻일 뿐, 제3자 제공에 대한 동의를 대가로 별도의 인센티브(incentive)를 부여하는 것도 허용되지 아니한다는 뜻은 아니다. 즉, 개인정보인 데이터도 계약상 반대급여 내지 대가가 될 수 있고,⁵⁸ 이 경우에 부당결부금지(Koppelungsverbot, 개인정보 보호법 제22조 제5항) 위반 문제는 생기지 아니하는 것이다. 또한 이러한 동의는 원칙적으로 제공의 대상인 제3자와 제공받는 자의 이용목적, 제공하는 개인정보의 항목, 보유 및 이용기간이 특정되고 동의를 거부할 수 있음이 명시된 것이어야 한다(개인정보 보호법 제18조 제3항, 제22조 제2항).⁵⁹ 특히 개인정보인 데이터의 거래가 문제되는 마케팅 목적 주소거래의 경우 그러한 사정을 명확히 알리고 동의를 받을 필요가 있다(개인정보 보호법 제22조 제4항).

다음 개인정보를 “통계작성 및 학술연구 등의 목적을 위하여 필요”하여 제공하는 경우에는 동의 없이도 제3자제공이 가능하다(개인정보 보호법 제18조 제2항 제4호). 통계작성은 통계 값을 도출하기 위한 것으로 결과에서 개인정보가 완전히 사라진다는 점에 특징이 있다. 통계작성이 공공기관 기타 통계법상 통계작성기관에 의한 것으로 제한되는가 아니면 영리기업의 마케팅 등 목적 통계작성도 이에 포함되는가 하는 점은 논란의 소지가 있는데, 유럽의 해석을

⁵⁸ 김진우, “대가로서의 디지털 개인정보 - 데이터의 개인정보보호법 및 계약법적 의의 -”, 비교사법 제24권 제4호(2017), 1521면 이하; Langhanke, Daten als Leistung, 2018, S. 76 ff.

⁵⁹ Specht(주 37), S. 41 f.

참조하여 후자도 포함된다는 견해가 있다.⁶⁰ 학술연구에 영리 목적의 연구가 포함되는가 하는 점은 더욱 격렬하게 다투어져, 영리기업이 지원한 연구도 안 된다는 견해, 영리기업이 지원한 것은 무방하나 공적 연구기관에서 수행되어야 한다는 견해 및 영리 목적의 연구도 포함된다는 견해가 주장되고 있고, 해석상으로는 정설(定說)이 없다.⁶¹ 이러한 사유를 들어 제3자 제공을 하는 경우에는 위와 같은 사정을 고려하여야 한다. 다른 한편, 이때 제3자제공은 “특정 개인을 알아볼 수 없는 형태”로 이루어져야 한다(개인정보 보호법 제18조 제2항 제4호 후단). 이는 이른바 가명처리(pseudonimization)를 해야 함을 의미한다. 가명처리된 정보도 개인정보이므로 이는 개인정보의 제3자 제공에 해당한다. 앞서 동의를 받은 경우에도 가명처리가 가능하면 가명처리하여야 한다.⁶²

나아가 개인정보를 제3자에게 제공할 때에는 그 제공환경, 가령 개인정보인 데이터 자체를 이전해줄지, 접속권한을 부여하여 외부에서 접속할 수 있게 해줄지, 개인정보처리자 내부의 지정된 공간에서 제한적 접속만을 허용할지도 검토하여야 한다. 제3자 제공 자체는 정당하다 하더라도 그로 인하여 위법한 침해 사고의 위험 또한 높아질 수 있기 때문이다.

개인정보를 제3자에게 제공하는 것은 대체로 개인정보 침해 위험을 높이는 일이므로, 신중한 검토가 필요하다. 개인정보 보호법 제31조는 개인정보처리자에게 개인정보의 처리에 관한 업무를 총괄해서 책임질 개인정보 보호책임자(Data Protection Officer)를 지정할 것을 요구하고 그의 업무를 열거하고 있다. 개인정보의 제3자 제공은 직접 열거되어 있지 아니하지만 그럼에도 불구하고 개인정보 보호책임자가 그 잠재적 위험 등을 검토하여야 할 사항이라고 봄이 옳다. 개인정보 보호책임자는 제공 가능성 및 제공 여부, 제공 방법을 검토하여 ‘개선조치’를 취하여야 한다. 물론 최종 결정권은 개인정보처리자에게 있고, 개인정보 보호책임자는 그에게 조언을 하거나 위임받은 범위 내의 권한을 행사할 수 있을 뿐이다.

나) 개인정보 처리동의를 흠결·철회와 데이터거래계약

개인정보의 처리에 대한 동의는 인격적 법익에 대한 법적행위(Rechtshandlung)로서 동의(Einwilligung)이자 단독행위이다. 이는 법률행위와 유사하나 원칙적으로 자유로이 철회할 수 있다는 데 특징이 있다. 개인정보 보호법도 처리정지요구권(제37조)을 규정하고 있고, 실무상으로도 철회가 가능하다고 보고 있다. 철회의 결과 당초 정보주체와 사이에 체결한

⁶⁰ 박광배·채성희·김현진(주 31), 175면 이하; 전용준, “공공데이터의 이용과 통계 및 학술연구 목적의 데이터 처리”에 한 연구 - 데이터의 안전한 이용의 관점에서 -, 정보법학 제23권 제1호(2019), 128면 이하.

⁶¹ 전용준(주 60), 132-133면. 또한 이동진(주 15), 258면.

⁶² 개인정보 보호법 제3조 제7항은 이를 ‘익명처리’라고 한다.

계약이 더는 이행될 수 없게 되었다면, 철회는 필연적으로 계약해지를 포함하게 된다. 정보주체는 계약을 언제든지 임의 해지할 수 있다. 다만 정보주체가 해지의사를 명확히 하지 아니하고 처리동의만 철회하겠다고 하는 경우에는 (해지까지 해야 한다는 의미에서) 처리정지요구를 거절할 수 있을 뿐이다(제37조 제2항 제4호).

정보주체가 제3자 제공에 동의하지 아니하였거나, 아예 처리 일체에 동의하지 아니하였고, 그 이외에 수집에 정당한 근거가 (더는) 존재하지 아니한다면, 개인정보처리자가 데이터제공자가 되어 데이터수령자와 체결한 데이터제공계약도 그 범위에서는 강행규정 위반으로 무효가 된다. 급여목적이 그 자체 불법인 제3자 제공을 지향하기 때문이다.⁶³ 다만 이는 데이터수령자가 계약이행을 구할 수 없다는 의미에서 그러한 것이고 (특히 데이터수령자가 선의인 경우) 손해배상도 구할 수 없다는 뜻은 아니다. 다수의 개인정보를 제공하는 경우 그중 일부에 동의의 흠결이 있을 가능성이 있으므로 이 경우의 책임 등을 명확히 하는 것이 바람직하다.⁶⁴

문제는 정보주체의 제3자 제공 동의에 터 잡아 데이터제공계약이 체결되고 이행되었는데, 이후 정보주체가 위 동의를 철회하는 경우이다. 이때에는 원칙적으로 추가 처리가 중단되어야 하고, 제3자 제공도 더는 불가능해진다. 두 가지 점에 주의하여야 한다. 첫째, 제3자 제공에 대한 동의는 정보주체와 데이터제공자 사이의 본래의 계약과 독립하여 이루어지므로, 제3자 제공에 대한 동의가 별도의 인센티브와 결부되고 그 자체 별도의 부수적 계약을 이루는 경우가 많다. 이때 제3자 제공 동의를 철회는 부수적 계약의 해지를 전제하게 마련이다.⁶⁵ 그러한 해지 및 청산이 불가능하거나 부적절한 경우에는 동의철회가 신의칙(민법 제2조 제1항)에 반하여 허용되지 아니할 수 있다.⁶⁶ 둘째, 제공된 데이터가 식별정보는 포함하지 아니하나 식별가능성까지 배제하지는 못하여 여전히 개인정보인 경우, 정보주체가 동의를 철회하면 데이터수령자의 추가처리를 정지시켜야 한다. 이때 데이터수령자에게 이전된 데이터 중 어느 데이터가 당해 정보주체에게 귀속되는 것인지 데이터제공자도 데이터수령자도 특정하기 곤란할 수 있다. 그러나 이 경우에도 법적으로는 여전히 개인정보이므로 처리정지의무가

⁶³ Langhanke(주 58), 147 ff.; Specht(주 37), S. 172 ff.

⁶⁴ 가능한 계약형성방식은 여러 가지가 있다. 가령 전체 데이터의 일정비율 내의 동의의 흠결에 대하여는 책임을 묻지 아니하기로 하거나, 건당 일정금액의 수수료/처리비용/ 손해배상 등을 약정하는 등.

⁶⁵ 앞서 본 바와 같이 이 경우 데이터제공은 반대급부에 해당한다. 데이터제공이 반대급부라 하더라도 인격권적 이익인 이상 계약의 구속력을 피하여 철회할 가능성이 있으나, 그러한 철회가 허용된다면 그때에는 거꾸로 계약이 해지되어야 하는 것이다. 개인정보 보호법 제22조 제2항 제4호도 참조.

⁶⁶ Buchner, Informatoinelle Selbstbestimmung im Privatrecht, 2006, S. 270 ff.

존재하지 아니한다고 할 수 없고, 그 결과 '구체적 손해'가 발생하면⁶⁷ 책임을 질 가능성이 있다.

따라서 개인정보인 데이터를 제공할 때에는 식별정보를 제외하는 경우에도 데이터제공계약상 정보주체의 동의철회 등으로 일부 데이터의 처리가 중단될 수 있다는 점을 명확히 하여야 하고, 그 경우, 특히 유상계약일 때에는 가급적, 계약청산에 관한 규정을 포함시켜야 한다. 또한 식별정보를 제외하는 경우에도 데이터제공자는 추후 데이터수령자에게 제공된 데이터 중 정보주체에게 귀속된 데이터를 식별할 수 있는 별도의 매칭테이블을 가지고, 그 매칭테이블에 터 잡아 필요하다면 제외할 데이터를 재식별해낼 수 있게 하여야 한다.

그 이외에 파생데이터의 귀속과 처리에 관하여도 동의철회의 가능성을 배려할 필요가 있다. 파생데이터가 여전히 식별가능성이 있어 개인정보인 경우 파생데이터에 대하여도 정보주체의 동의철회의 효력이 미친다.⁶⁸ 파생데이터가 개인정보에 해당하는 경우에는 파생데이터의 사실상 지배의 귀속에 관하여 어떻게 약정하였든 동의철회를 고려한 추가 및 처리정지, 삭제청구 가능성을 확보해둘 필요가 있다.

2) 비식별화·익명화된 개인정보의 거래

가) 재식별위험에 대한 대응

개인정보가 비식별화·익명화되면 비개인정보로 전환되므로 원칙적으로 개인정보 보호법의 적용을 받지 아니한다. 그러나 식별가능성의 판단기준 자체에 모호성이 있을 뿐 아니라, 식별가능성을 제거하였다 하더라도 결합정보의 증가, 결합정보에의 접근가능성의 확대, 분석기법의 고도화에 따라 시간이 지나면서 재식별가능성이 생길 수도 있다. 그러므로 개인정보를 비식별화·익명화한 경우에도 여전히 재식별위험을 차단하기 위한 추가적인 조치가 필요하다. 이는 개인정보를 가명화하여 제공하는 경우에도 마찬가지이다. 국무조정실 등 관계부처 합동으로 2016. 6. 발간한 「개인정보 비식별 조치 가이드라인 - 비식별 조치 기준 및 지원·관리체계 안내」는 제3자제공시 계약서에 재식별금지 및 재식별위험이 높아지는 경우 처리중단과 통지의무 등을 포함시키고, 그에 대한 책임을 명시할 것을 권고하고 있다.⁶⁹ 재식별위험 증가로 인하여 더는 데이터제공 및 이용이 불가능해진 경우 그 책임 소재에 따라 계약을 어떻게 청산할 것인지에 대하여도 규정할 필요가 있다.

⁶⁷ 구체적 손해가 어떠한 경우에 인정되는지에 대하여는 여러 모호한 점이 있으나, 식별성이 없는 상태가 유지되는 한 식별가능성만으로는 구체적 손해가 부정될 여지가 있다. 이동진(주 15), 269면 이하.

⁶⁸ 박광배·채성희·김현진(주 31), 175면 이하.

⁶⁹ 국무조정실 등, 「개인정보 비식별 조치 가이드라인 - 비식별 조치 기준 및 지원·관리체계 안내」, 2016, 16면.

나) 결합정보의 생성과 데이터거래

비개인정보인 데이터의 결합은 데이터제공자가 하든 데이터수령자가 하든 법적으로 특별한 문제를 야기하지 아니한다. 데이터제공자가 하는 경우에는 그 범위에서 데이터창출이 되고, 도급의 성격을 가질 뿐이다. 반면 개인정보인 데이터의 결합은 그 자체 개인정보의 처리일 뿐 아니라 정보량을 증가시켜 식별가능성을 더욱 높인다. 정보주체의 동의가 있다면 어떻든 법적 문제는 없으나, 정보주체의 동의 없이 이를 수행하는 데는 여러 법적 문제가 생긴다. 그러나 다른 한편 서로 다른 수준에서 수집된 데이터를 결합시켜 연결관계를 파악하는 데서 데이터의 가치가 극대화되는 경우가 많다.

앞서 본 「개인정보 비식별 조치 가이드라인 - 비식별 조치 기준 및 지원·관리체계 안내」는 비식별 데이터 결합을 두 단계로 진행하도록 하고 있다. 첫 번째 단계는 각 데이터제공자가 각자 비식별화를 진행하는 것이고, 두 번째 단계는 비식별화된 두 데이터셋을 결합하는 것이다. 문제는 두 데이터셋을 결합하기 위해서는 각 데이터셋의 항을 대응시킬 매칭키가 필요하고, 각자의 원본데이터는 식별자를 포함하고 있으므로 결국 매칭키는 식별자와 매칭키 사이의 연결을 포함하게 되는데, 각 데이터제공자 중 어느 한 쪽이 결합을 하게 되는 경우 비식별데이터와 매칭키 양쪽을 다 갖고 있어 그에게는 개인정보가 될 가능성이 높고,⁷⁰ 결합 자체가 개인정보의 처리가 되어 동의 없이는 합법성을 확보하기 어려워진다는 데 있다.⁷¹ 그러므로 (신뢰할 수 있는) 제3자(Trusted Third Party)를 개입시켜 각 데이터제공자는 동일 알고리즘을 적용하여 복호화가 불가능한 임시대체키를 생성하고 비식별화를 진행한 다음, 제3자에게 비식별화된 데이터를 제공하여 제3자가 결합을 수행하고 임시대체키를 완전 삭제하는 방식을 권고하고 있다. 비식별조치에 대한 적정성 평가는 두 단계 모두에서 하여야 한다.

그러나 이러한 방식에는, 그 표현에도 불구하고, 가명식별자에 의한 결합으로 평가될 위험이 존재한다. 각 데이터제공자가 데이터셋을 ‘비식별화’하면서 식별자만 매칭키로 대체하고 각 데이터의 유일성을 제거하지 아니한 채⁷² 데이터를 제공한다면, 이 데이터는 가명데이터이고, 각 데이터제공자는 원데이터와 매칭키로 전환된 데이터를 모두 가진 채 신뢰할 수 있는 제3자에게 이를 제공하므로, 위 제공 자체가 개인정보의 제3자 제공이 된다. 즉, 이 단계의 비식별화를 법적인 의미의 비식별화라고 할 수 없는 것이다. 이론적으로는 데이터셋을 ‘비식별화’할 때 식별자를 매칭키로 대체하는 것이 아니라 각

⁷⁰ 앞서 본 바와 같이 약학정보원 사건의 항소심 판결은 이 사실만으로도 개인정보가 된다는 입장을 취하였다.

⁷¹ 이것이 현행법 위반이 될 수 있다는 점은, 국무조정실 등(주 69), 18면.

⁷² 그렇게 하여야 정보량이 최대한으로 유지되고, 결합이 완전하게 수행된다.

데이터항의 유일성까지 제거하여 원데이터와 비식별데이터를 동시에 갖고 있어도 양자를 일대일 대응시킬 수 없게 하는 것이 가능하나,⁷³ 아직까지 잘 쓰이고 있지는 아니하다. 비식별조치 가이드라인은 법규범이 아니므로, 유일성 있는 매칭키와 신뢰할 수 있는 제3자를 활용하는 형태의 데이터 결합은, 별도의 법적 근거가 없는 한,⁷⁴ 불법으로 평가될 위험이 남아 있다. 물론 통계처리 및 학술연구 등의 목적의 개인정보이용 및 제3자 제공을 허용하는 개인정보 보호법 제18조 제2항 제4호를 근거로 이 정도의 가명키 대체도 적법하다고 주장할 가능성은 있다.⁷⁵

개인정보와 비개인정보의 구분은 결합정보의 보유 및 접근 가능성 등 여러 사정에 달려있는 상대적 개념이다. 그러므로 결합데이터를 생성하여 제공하는 경우 각 관계자들이 자신 또는 상대방의 결합정보에 접근할 가능성을 차단하여야 처리의 적법성이 확보된다. 특히 익명화의 정도가 높지 아니하여 결합정보가 있으면 재식별위험이 생기는 경우 각각의 데이터제공자와 결합을 수행할 신뢰받는 제3자 사이에 매칭키가 포함된 비식별데이터의 교환 외에 재식별위험을 증가시키는 일체의 접촉·정보교환 등을 차단하고, 데이터수령자 측도 재식별위험이 발생 또는 증가하는 경우 추가 활용을 중단하고 폐기 또는 재처리하는 등 조치를 취하도록 의무화할 필요가 있다. 기본적으로 한 명의 데이터제공자가 존재하는 경우와 같으나, 데이터결합의 경우 세 명 이상이 관여하게 마련이므로, 데이터계약상 각자의 책임을 명확히 하여야 한다.

다) 투명성과 가치공유, 대가결정

개인정보의 공유는 당초 정보주체가 제공한 정보를 제3자와 공유하는 것으로, 그 필요성과 가치에도 불구하고 정보주체의 위험을 증가시킬 가능성을 내포하고 있다. 이는 개인정보를 비식별화·익명화하여 제공하는 경우에도 마찬가지이다. 그러므로 개인정보를 비식별화·익명화하여 제공하는 경우에도, 법적인

⁷³ 관련 문제는 이원석, “익명화 데이터의 익명 결합 방법”, 전자금융과 금융보안 제15호(2019), 88면 이하 참조. 같은 문헌은 이러한 기법을 사용하는 경우 (바로 위에서 지정한 바와 같은 이유에서) 신뢰할 수 있는 제3자(TTP) 개입 없이 직접 결합하는 것이 가능해진다고 설명한다. 비개인정보이기 때문이라는 것이다.

⁷⁴ 국회에 계류 중인 개인정보 보호법 개정안은 이 부분의 입법화를 포함하고 있으나, 개정 여부는 현 시점에서는 불투명하다.

⁷⁵ 다만 이 규정은 GDPR 제89조와 달리 일체의 ‘처리’에 적용되는 예외가 아니라 ‘이용’과 ‘제공’에 대한 예외로 되어 있어, ‘가공, 연계, 연동’(개인정보 보호법 제2조 제2호 참조)에도 적용되는지 다툼이 생길 소지가 있다.

요구사항은 아니나, 일정한 투명성과 가치공유가 필요하다.⁷⁶ 비식별화·익명화된 데이터도 결국 정보주체가 제공한 개인정보에서 유래한 것이므로, 공중(公衆)의 신뢰를 잃는다면 장기적으로는 그 원재료인 개인정보를 확보하기 곤란해질 가능성이 있기 때문이다. 이는 개인정보인 데이터의 활용에 있어 그 성과가 간접적인 형태라라도 공익 또는 정보주체의 이익에 기여하는 것이 바람직함을 뜻한다.

이 점은 개인정보를 비식별화·익명화한 데이터를 제공하는 경우 그 대가의 결정에 대하여도 시사하는 바가 있다. 데이터제공계약의 원재료가 개인정보인 경우 이를 비식별화·익명화하여 제공한다 하더라도 데이터제공자가 그로부터 과도하게 상업적 이익을 추구한다는 인상을 주는 경우 '내 정보를 가지고 기업이 돈을 번다'는 대중적 반응을 초래할 위험이 있다는 점에 유의하여야 하는 것이다.

4. 나가며

데이터제공과 재사용의 활성화가 데이터경제와 사회의 진전에 긴요하다는 데는 이론(異論)이 없다. 그러나 그 과정에서 제기되는 여러 쟁점에 대한 법적 대응은 충분하지 아니하다. 이는 데이터 자체의 법적 취급이 아직까지 분명하지 아니하다는 점과 특히 개인정보의 경우 그 보호의 필요성과 지속적인 긴장관계에 놓인다는 점에서 비롯한다. 이러한 상황을 극복하기 위해서는 높은 신뢰성과 경험을 가진 데이터거래 주도와 거래를 가능하게 하는 제3자의 보증, 가령 데이터 가치평가 기관들 및 다양한 데이터공유플랫폼이 나타날 필요가 있다. 데이터거래의 표준과 상호운용성 확보, 보안성 강화도 중요하다.

이 글에서는 그중 특히 데이터거래의 고려요소와 그 결과로 나타나는 거래의 구조·내용을 검토하였다. 데이터거래의 많은 부분이 거래과정과 계약형성에 달려 있다는 점에서, 단계별로 어떠한 법적 쟁점이 있고, 법적 또는 사실상 어떠한 점을 고려하여야 하는지를 정리하였다. 좀 더 대담한 접근을 위해서는 법 개정과 시장조성활동이 있어야 하나, 이 부분은 데이터공유플랫폼의 법적 구조 및 쟁점과 함께 별도의 작업을 기약하기로 한다.

⁷⁶ 유럽연합은 데이터 공유에 관한 핵심 원칙(key principles)으로, 투명성(transparency), 가치창출의 공유(shared value creation), 상호 상업적 이익의 존중(respect for each other's commercial interests), 왜곡되지 아니한 경쟁의 보장(ensure undistorted competition), 데이터 고착의 최소화(minimized data lock-in)를 든다. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS "Towards a common European data space" (SWD(2018) 125 final), 2018, p. 10. 영국의 ICO는 개인정보 공유에 대하여 공정성과 투명성, 보안, 거버넌스를 고려요소로 든다. ICO, Data sharing code of practice, 2011, pp. 17 ff.

1. 데이터거래의 준비

데이터거래의 의의

- 데이터는 유용한 정보를 전자기적인 방식으로 읽어 들일 수 있게 기입한 것으로 그 매체와 내용은 가리지 아니한다.
- 오늘날 기업·단체·공공기관·개인의 일상적인 활동과정에서 다량의 데이터가 매 순간 수집·축적되고 있다. 이를 공유(share)하거나 재사용(re-use)하는 것은 데이터보유자뿐만 아니라 거래 상대방, 일반 공중에게도 경제·사회적으로 큰 의미가 있다.

데이터관리계획

- 데이터보유자는 보유데이터의 종류와 수량, 품질, 가치 및 위험성을 평가하고 적합한 조치를 유지하여야 한다. 이러한 조치에는 암호화 등 보안기술과 접근(log)기록·필요최소한의 처리 등 관리방침이 포함된다.
- 데이터보유자는 보유데이터에 대하여 필요한 수준의 보안성과 상호운용성(interoperability)을 확보하여야 한다.
- 데이터보유자는 보유데이터의 취득경위와 그 법적 근거, 보유데이터에 대한 제3자의 권리보유 여부를 미리 확인·평가하여야 한다.

데이터거래시장에서의 참여

- 데이터거래는 여러 가지로 분류할 수 있다. 오픈 데이터 접근, (일대일) 데이터거래, (폐쇄된) 데이터공유플랫폼 내의 교환으로도 나눌 수 있고, 데이터제공, 데이터창출, 데이터공유로 나눌 수도 있다.
- 오픈 데이터 접근은 비교적 제한 없이 데이터를 공개하여 재사용할 수 있게 해주는 것을 말한다. 이러한 접근을 취하는 경우 개방할 수 있는 데이터를 선별하여 보유자가 스스로 운영하는 플랫폼에서 일정요건만 갖추면 데이터에 접속하여 이용할 수 있게 하거나 오픈데이터플랫폼에 업로드 하는 방식으로 공유하게 마련이다. 데이터는 한 번 공개되면 다시 이용을 제한하기 어려워 공개로 인하여 당해 데이터보유자 또는 제3자에게 손해가 생기는 경우 회복할 수 없으므로 현실적으로 그 적용범위가 제한된다.
- (폐쇄된) 데이터공유플랫폼 내의 교환은 플랫폼서비스제공자가 데이터공유에 관한 규약을 제정하고 참여자들을 모집하여 참여자들 사이에서 데이터를 공유하는 것을 말한다. 이때 참여자들은 직접 데이터를 주고받은 상대방이 아니라 하더라도 당해 플랫폼의 데이터공유에 관한 규약의 구속을 받는다는 장점이 있다. 데이터보유자는 당해 플랫폼이 제공하는 조건에 따라 제공 가능한 데이터를 선별하여 데이터공유플랫폼을 통하여 공유한다. 플랫폼규약 중에는 공유데이터로 분류된 이상 추가조건 없이 자유롭게 이용하는 것을 허용하는 것이 있다. 각 당사자가 독자적인 제공조건을 설정하는 데 제약이

부과되는 경우가 많고 맞춤형 데이터와도 다소 거리가 있다. 이 방법도 그 적용범위에 다소 제한이 있다.

- (일대일) 데이터거래는 잠재적인 데이터수령자가 잠재적인 데이터제공자에게 보유데이터의 제공을 요청하거나 반대로 잠재적인 데이터제공자가 잠재적인 데이터수령자에게 제공을 제안함으로써 이루어질 수 있다. 데이터브로커가 중간에 개입하는 경우도 상정할 수 있다. (일대일) 데이터거래는 다양한 거래의 기초를 이루고 여러 형성가능성이 있어, 위 두 유형의 기본형이라고 할 수 있다.
- 이미 존재하는 데이터를 제공하는 것이 아니라 특정 필요에 따라 데이터를 생성하여 제공하는 경우를 생각할 수 있는데, 이를 데이터창출이라고 한다. 복수의 데이터 사이의 결합을 통하여 새로운 결합데이터를 생성하는 것도 이에 포함된다. 데이터가 제공자의 사정과 수령자의 필요에 따라 가공되어 제공되어야 하는 경우가 많은 이상 데이터제공과 데이터창출의 경계는 유동적이다.
- 이 가이드라인은 주로 공개여부, 플랫폼참여여부 및 범위만 문제되는 위 두 유형보다는 거래 당사자 사이의 구체적 거래형태와 계약내용 결정이 문제되는 (일대일)데이터거래에 초점을 맞춘다. 다만, 그와 관계되는 범위에서 데이터창출도 다룬다.

2. 데이터제공여부의 결정

데이터의 구별

- 데이터는 여러 방법으로 구분할 수 있다. 그중 특히 중요한 것은 구조화된 데이터와 구조화되지 아니한 데이터, 개인정보인 데이터와 비개인정보인 데이터, 영업비밀인 데이터와 영업비밀로 보호되지 아니하는 데이터의 구분이다.
- 구조화된 데이터는 활용이 편리하나 개인정보인 경우에는 개인정보파일을 이루어 규제의 대상이 될 가능성이 높고, 개인정보인지 여부와 관계없이 구조화정도에 따라 데이터베이스제작자로서의 권리가 인정될 가능성이 있다.
- 개인정보인 데이터는 개인정보 보호법 등의 적용을 받는 반면, 개인정보가 아닌 데이터는 그렇지 아니하다는 차이가 있다.
- 영업비밀로 보호되는 데이터는 별도의 권리가 없어도 부정경쟁방지 및 영업비밀보호 등에 관한 법률에 따라 금지청구, 손해배상, 형사처벌 등으로 보호받는 반면, 영업비밀로 보호되지 아니하는 데이터는 그러한 보호를 받지 못한다. 영업비밀로 보호받으려면 데이터를 비밀로 관리하여야 한다. 어느 정도의 비밀관리가 있어야 영업비밀로 보호받을 수 있는지에 대하여는 관련 분야 전문가의 의견을 구할 필요가 있다.

개인정보인 데이터와 비개인정보인 데이터의 구별

- 그중 특히 중요한 것은 개인정보인 데이터와 비개인정보인 데이터의 구별이다.

전자에 대하여는 개인정보 보호법상 데이터거래에 직접적인 제약이 있을 뿐 아니라 데이터가 귀속되는 살아 있는 개인, 즉 정보주체가 품는 불안감, 수익배분요구 등에 대하여 대응할 사실상의 필요도 있다.

- 개인정보는 살아 있는 개인에 대한 정보로서 그 정보만으로 또는 다른 정보(결합정보)와 쉽게 결합하여 그 정보주체가 누구인지 식별할 수 있는 것을 말하고, 비개인정보는 개인정보가 아닌 정보 일체를 말한다. 어느 한 사람에 대하여 인종, 성별, 나이, 출생지, 직업, 건강상태 등 다수의 정보가 있으면, 당해 정보 안에 그 사람의 이름과 주소 등 직접 정보주체를 식별하는 정보(식별정보)가 없다 하더라도 다른 정보와 결합하여 정보주체를 식별해낼 수 있는 것이 보통이므로, 관건은 '쉽게 결합하여 식별할 수 있는'의 의미 내지 그 정도이다. 이에 관하여는 존재하거나 상정할 수 있는 모든 결합정보와 분석기법을 고려할 필요는 없다는 데 의견이 모아지고 있으나, 정보주체를 식별하는 것이 얼마나 어려워 비개인정보가 되는지 그 정도에 대하여는 논란이 분분하고, 사법기관의 입장도 확고하다고 할 수 없다. 따라서 살아 있는 특정 개인에게 귀속될 가능성이 있는 정보를 다루는 경우에는 그 정보가 개인정보인지 내지 개인정보로 분류될 가능성(위험)이 어느 정도인지에 대하여 관련 분야 전문가의 의견을 구할 필요가 있다.
- 개인정보의 식별가능성을 더는 개인정보로 분류되지 아니할 정도까지 낮추는 조치를 비식별화(de-identification)·익명화(anonymization)라고 한다. 개인정보처리자가 이 조치를 취하면 그 정보는 더는 개인정보가 아니고, 따라서 개인정보 보호법상의 규제를 피할 수 있다. 다만 기술적인 방법으로 데이터의 식별가능성을 완전히 제거하는 것은 종종 정보량의 중대한 손실을 수반하여 데이터의 가치를 떨어뜨린다는 문제가 있다. 따라서 비식별화·익명화는 데이터의 활용목적에 필요한 수준의 정보만을, 그러나 필요한 수준의 정보는 가급적 많이, 그리고 정확하게 포함시킬 수 있는 기법을 쓰고, 그럼에도 불구하고 남은 위험은 관리적 조치, 즉 데이터거버넌스로 보완하는 방식으로 이루어지게 된다. 비식별화·익명화되었다 하더라도 결합정보가 생성 또는 취득되거나 데이터 분석기법이 고도화되면 일단 비식별화·익명화된 데이터가 다시 식별가능성을 갖게 될 수 있는데, 이때에는 위 데이터는 비개인정보에서 개인정보로 자동으로 전환된다. 그러므로 비식별화·익명화된 데이터를 비개인정보로 취급하는 때에는 재식별위험의 증가 여부를 지속적으로 감시하고 대응책도 데이터거래계약 단계에서부터 준비해둘 필요가 있다.
- 종종 오해되고 있는 것과 달리 식별정보를 임의의 번호, 기호, 가명으로 대체하는 가명화(pseudonimization)는 개인정보를 비개인정보로 바꾸지 아니한다. 그러나 가명화를 전제로 데이터를 활용할 수 있는 경우가 있을 뿐 아니라 가명화는 굳이 식별정보까지 필요하지 아니한 여러 데이터 처리에서 개인정보 오·남용 위험을 줄여주는 장치로 적극 권장된다.

비개인정보인 데이터 제공여부의 결정

- 비개인정보인 데이터 제공여부에 관한 사실상의 고려요소로는 다음을 들 수 있다.
 - 당해 데이터가 저작권 등 제3자의 권리의 대상인가
 - 당해 데이터가 기업 경쟁력의 중요 원천인가
 - 당해 데이터가 비밀로 유지되고 있는 정보인가
 - 당해 데이터의 가치와 양은 어느 정도인가
 - 당해 데이터를 통하여 받을 수 있는 대가는 무엇인가
 - 데이터거래의 상대방의 신용과 보안성을 신뢰할 수 있는가
- 데이터의 내용이 저작권·특허권 등 제3자의 권리의 대상이고, 데이터의 제공이 반드시 또는 매우 높은 확률로 제3자의 권리침해를 수반 내지 야기하는 경우에는 데이터 제공에 신중하여야 한다. 제3자 권리침해가 데이터수령자에 의하여 이루어진다 하더라도 데이터제공자도 공동으로 법적 책임을 질 가능성이 있다.
- 당해 데이터가 기업 경쟁력의 중요 원천이고 다른 사람이 그 정도의 데이터를 집적할 수 없다면 데이터 제공에 신중할 필요가 있다. 데이터는 현행법상 배타적 재산권의 대상이 아니어서 일단 공개·유통되면 그 회수가 사실상로는 물론 법적으로도 곤란할 가능성이 높으므로, 비밀로 유지되고 있는 데이터일수록, 그리고 상대방의 신용과 보안성을 신뢰할 수 없을수록 데이터 제공에 신중할 필요가 있다. 데이터 제공으로 받을 수 있는 대가가 충분한지 또한 고려되어야 한다.
- 그러나 다음과 같은 측면도 고려할 필요가 있다.
 - 첫째, 데이터를 통하여 시장지배력을 행사하는 기업의 경우 당해 데이터를 합리적인 조건으로 제공하지 아니하는 것이 시장지배력의 남용으로 평가되어 독점규제 및 공정거래에 관한 법률상 제재대상이 될 가능성이 있다.
 - 둘째, 데이터를 통한 시장지배력을 이용하여 데이터 거래 상대방에게 부당한 거래조건을 부과하는 경우, 가령 데이터 거래 상대방의 다른 데이터의 제공을 요구하거나 경쟁자와의 거래를 금지하는 경우 불공정거래행위로 평가되어 독점규제 및 공정거래에 관한 법률상 제재대상이 될 가능성이 있다.
 - 셋째, 법률위반이 아니라 하더라도 특정 영역의 데이터를 사실상 독점하는 기업이 공유 및 재사용에 소극적인 경우 데이터 라이프 사이클(life-cycle)을 위축시키는 행위로 비난의 대상이 될 수 있다.
- 데이터거래의 위험 중 상당부분은 적절한 거래형성을 통하여 대응할 수 있으므로 위험이 있다는 이유만으로 거래에 소극적일 필요는 없고, 전향적인 태도를 취할 필요가 있다.

개인정보인 또는 개인정보였던 데이터 제공여부의 결정

- 개인정보인 데이터의 제공은 개인정보 보호법상 제3자 제공의 요건을 갖추어야 한다. 즉, 정보주체에게 제3자 제공의 구체적인 내용에

- 대하여 설명하고 동의를 받았거나 가명화한 형태로 제공하는 것으로 '통계목적·학술목적 등'을 위한 경우여야 한다. 통계목적·학술목적 등의 제3자제공의 예외의 구체적 요건, 특히 기업이 상업적 목적으로 수행하는 통계조사나 연구가 이 요건을 충족하는지에 대하여는 상당한 논란이 있고, 아직 확립된 정설(定說)이 없으므로, 관련 분야 전문가의 의견을 구할 필요가 있다.
- 개인정보가 비식별화·익명화된 경우에는 그 제공에 별다른 법적 제한이 없다. 그러나 비식별화·익명화가 제대로 이루어졌는지 여부에 대하여는 언제나 불명확성이 있으므로, 2016년 국무조정실 등이 발표한 「개인정보 비식별조치 가이드라인」이 권고하는 바와 같이 외부 전문가 등의 평가를 거치는 것이 바람직하고, 책임위험을 줄여준다.
- 비식별화·익명화로 비개인정보가 되었는지 여부는 구체적 맥락의 영향을 받게 마련이다. 데이터제공자에 대하여 비개인정보가 되었다 하더라도 (따라서 별도의 분석목적으로 이를 활용할 수 있다 하더라도) 데이터수령자에게는 개인정보일 수 있다. 그러므로 비식별화·익명화에 근거하여 개인정보였던 데이터를 제공하는 경우에는 데이터수령자가 보유하는 결합정보와 그의 분석능력을 검토하여 개인정보 여부를 재평가하여야 한다.
- 그러나 개인정보 및 비식별화·익명화된 정보의 제공에는 그 정보주체가 드러나 개인정보자기결정권이 침해될 위험이 수반될 수밖에 없으므로, 데이터 제공으로 인하여 증가하는 위험이 어느 정도인지를 고려하여야 한다. 가령 데이터수령자의 신용과 보안능력을 신뢰할 수 없거나, 제공되는 데이터가 특히 민감한 내용을 포함한다거나 상대적으로 재식별위험이 높다는 등의 사정을 신중하게 고려할 필요가 있다.
- 개인정보인 또는 개인정보였던 데이터의 제공이 법적으로 가능하다 하더라도 개인정보의 상업적 이용, 특히 이를 통한 데이터보유자의 수익은 정보가 귀속되는 정보주체의 불안을 야기하거나 수익배분 요구에 직면할 수 있고, 장기적으로는 지속적인 개인정보를 저해하여 본래의 사업을 위협할 수 있다. 개인정보인 또는 개인정보였던 데이터 제공과 관련하여서는 특히 이용목적(공적 성격을 갖는지 영리적 성격이 강한지)과 결과의 직·간접적 공유(가령 간접적으로라도 정보주체 등의 편익에 기여한다든지) 등을 고려할 필요가 있다.
- 개인정보인 또는 개인정보였던 데이터의 제공은 전 과정에 걸쳐 개인정보 보호법상 개인정보 보호책임자의 검토를 거쳐야 한다.

국내제공과 국외제공의 구별

- 데이터는, 특히 가상공간에 저장되고 가상공간상에서 전송되는 경우, 소재지가 분명하지 아니하다. 그러나 데이터보유자와 데이터수령자에게는 소재지가 있고, 이것이 각종 법집행이 준거가 된다. 그런데 데이터수령자가 국외에 소재하는 경우 그에 대한 법집행이 사실상 어렵다. 데이터의 국외제공은

특별히 큰 위험을 내포한다. 특히 개인정보인 또는 개인정보였던 데이터의 국외제공은 개인정보 보호법상의 제한을 받는다.

3. 데이터제공관계의 형성

데이터거래계약과 데이터제공관계의 기능

- 데이터는 그 자체 소유권의 대상이 아니고, 경우에 따라 부정경쟁방지 및 영업비밀보호 등에 관한 법률과 불법행위법, 형법에 의하여 부분적 보호를 받을 뿐이다.
- 데이터는 복제가 용이하고 일단 공개되면 비가역적인 성질을 가지는 것이 보통인데, 그 결과 데이터제공자가 제3자에 대하여 법적 책임을 지는 일이 생길 수 있다. 즉, 데이터 법은 여러 측면에서 전형적인 위험규제의 성격을 포함한다.
- 데이터에 관한 법적 규율의 공백은 계약으로 대응하여야 하고, 데이터와 관련된 위험에 대한 대응은 데이터제공자·데이터수령자의 데이터거버넌스 구축으로 대응하여야 한다. 이는 데이터제공관계의 구체적 형성·유지를 요구한다.

데이터거래계약의 구체적 내용(1): 준거법

- 데이터는 물리적 소재지를 논하기 어렵고, 배타적 재산권의 대상도 아니다. 데이터 침해에 대하여는 불법행위지법과 결과발생지법이 준거법이 될 가능성이 높으나, 법적 불확실성을 감소시키기 위하여 준거법에 관한 합의를 하는 것이 바람직하다. 이 가이드라인은 한국 법을 준거법으로 하는 합의를 하는 것을 전제로 설명한다.

데이터거래계약의 구체적 내용(2): 데이터의 귀속과 데이터거래의 유형

- 데이터공유 내지 이용허락
데이터거래 중 가장 보편적인 유형은 공유(sharing) 내지 이용허락(licensing)이다. 이 둘은 같은 개념으로, 데이터제공자가 데이터를 여전히 보유하고 대개 이용까지 하면서, 데이터수령자에게 데이터의 복사본을 제공하고 그 이용을 허락하는 것을 내용으로 한다. 그러나 데이터는 배타적 재산권의 대상이 아니어서 법적인 의미의 공유 대상이 될 수 없고, 이용허락도 지식재산권에서와 같은 의미를 가지지 아니하므로, 단지 공유 내지 이용허락이라고 하기 보다는 정의규정을 두어 그 구체적 내용을 설명하는 것이 바람직하다. 이때 다음과 같은 점을 구체적으로 규정하여야 한다.
 - 데이터수령자의 이용목적과 범위, 제3자(재)제공의 허부와 범위
 - 데이터제공자의 (동시/경합하는) 이용가능성과 그 범위
 - 데이터공유/이용허락의 기간과 해지사유

- 데이터공유/이용허락의 종료 후 조치(데이터 삭제 및 확인 등)
 - 데이터에 대한 종국적 권리 내지 처분권은 데이터제공자에게 유보된다는 점
- 데이터양도
데이터를 양도하는 경우 데이터제공자는 데이터수령자에게 데이터의 유일한 원본을 인도하거나 복사본을 제공 내지는 전송한 다음 데이터제공자에게 남은 데이터 복사본을 삭제한다. 이때 데이터에 대한 종국적 권리 내지 처분권이 데이터수령자에게 이전한다는 점과 데이터제공자는 더는 그러한 권리를 갖지 아니한다는 점을 명확히 규정할 필요가 있다. 성질상 데이터수령자의 이용목적과 범위, 제3자 제공에 대한 제한을 규정하는 일은 생각하기 어렵다. 그러나 제한된 범위에서 데이터제공자의 이용을 허락할 수는 있다.
 - 데이터창출
데이터를 새로 만들어주기로 하는 경우에는 도급계약이 된다. 그러나 데이터는 동시이용이 가능하므로 데이터거래계약은 주문에 따라 공급한 데이터가 데이터수령자에게 전적으로 귀속하는지, 아니면 데이터제공자에게 귀속하고 데이터수령자는 제한된 이용권만 갖는지를 분명히 하여야 한다. 즉, 도급으로서의 성격을 제외하면 앞의 두 유형의 거래의 고려요소가 여기에서도 그대로 고려되어야 한다.

데이터거래계약의 구체적 내용(3): 제공대상과 대가의 결정

- 제공의 대상이 되는 데이터를 명확히 규정하여야 한다. 데이터가 이미 존재하는 경우에는 데이터의 양과 질을 검토한 뒤 거래하는 것이 바람직하고, 그렇지 아니한 경우에는 이용목적 기타의 방법으로 요구되는 데이터의 양과 질을 특정하고, 그 내용이 데이터거래계약의 일부로 편입되도록 하여야 한다.
- 제공 대상이 되는 데이터의 포맷과 제공방법을 특정할 필요가 있다.
- 제공 대상이 되는 데이터의 표현 또는 내용에 데이터제공자 또는 제3자의 저작권·특허권 등 권리가 설정되어 있는 경우, 저작권·특허권 등의 권리에 관한 규율을 포함할 필요가 있다. 가령 저작권·특허권침해 등에 해당하는 이용이 배제되어 있는 경우에는 이용에 관한 규정에서 이를 명시하여야 하고, 그러한 이용이 포함되는 경우에는 저작물이나 특허의 이용허락 또는 양도가 포함되는지 여부를 분명히 하는 것이 바람직하다.
- 데이터의 금전적 가치 평가에는 여러 문제가 있다. 금전적 평가가 곤란하다고 판단되는 경우 데이터를 서로 교환하거나 데이터수령자의 이익을 배분받거나 데이터제공의 성과인 파생데이터에 대하여 일정한 권한을 취득하는 대안을 생각할 수 있다. 다만, 개인정보인 또는 개인정보였던 데이터를 금전적 대가를 받고 제공하는 경우 수익배분요구에 직면할 가능성이 있고, 데이터를 지배하는 사업자가 거래 상대방으로부터 금전이나 수익 이외에 아무 관련이 없는 별개의 데이터를 요구하거나 경쟁자 등과의 거래를 제한하는 경우에는 불공정거래행위로 평가될 위험이 있다는 점에 주의하여야 한다.

데이터거래계약의 구체적 내용(4): 파생데이터의 귀속

- 데이터수령자가 제공받은 데이터를 이용하는 경우 그 결과 새로운 데이터가 생성될 가능성이 높다. 데이터에 대한 배타적 재산권이 존재하지 아니하므로, 데이터수령자가 생성한 데이터의 이용권한도 데이터거래계약에서 규율하지 아니하면 법적으로 불확실한 상태가 된다. 그러므로 어떠한 파생데이터가 생성, 축적될 수 있는지 합의하고, 그에 대한 종국적 권한 내지 처분권의 귀속과 이용권한의 부여를 규정하여야 한다. 구체적 분배에 관하여는 여러 가지 방법이 있다. 데이터제공자와 데이터수령자 공동으로 보유할 수도 있고, 어느 일방에게 귀속시키되 일정한 범위에서 타방에게 이용권한을 부여하는 것도 가능하다.

데이터거래계약의 구체적 내용(5): 제공데이터에 대한 보증/비보증

- 데이터제공자가 제공데이터를 제공할 권한이 있다는 점에 대하여 약정을 하지 아니하는 경우 거래유형에 따라 매매법, 임대차법 등의 담보책임 규정이 적용된다. 가령 데이터양도의 경우 민법 제570조 이하에 따른 책임을 질 수 있다. 그러나 이에 관하여 명문의 규정을 두는 것이 바람직함은 물론이다. 가령 데이터에 대하여 제3자의 저작권·특허권 등의 대상이 아님을 보증하거나, 개인정보인 데이터에 대하여 제3자 제공에 대한 동의 등 적법요건을 갖추었다는 점을 보증하거나, 개인정보였던 데이터에 대하여 비식별화·익명화를 거쳤음을 보증하는 것이다. 물론 반대로 보증하지 아니함을 명시하거나, 전체 거래 대상 데이터 중 일정 비율까지는 제3자의 권리에도 불구하고 데이터수령자에게 어떠한 구제도 인정하지 아니하나 일정 비율을 넘으면 책임을 물을 수 있다고 규정할 수도 있다. 전면적 또는 제한적 보증을 제공하는 경우, 그 위반에 대한 정산방법을 데이터거래계약에 규정할 수도 있다. 가령 데이터 단위 당 일정금액의 반환을 약정할 수 있다.
- 제공데이터가 개인정보이거나 비식별화·익명화된 정보인 경우 제3자 제공이 적법하였다 하더라도 정보주체가 동의를 철회하거나 재식별위험이 생겨 더는 제공할 수 없게 되는 일이 생길 수 있다. 데이터거래계약은 한편으로는 각 데이터에 소유권을 부여하는 등으로 데이터제공자가 처리정지의 대상이 된 데이터에 공급할 가능성을 확보하고 데이터제공자의 요청이 있는 경우 처리정지에 데이터수령자가 협력할 의무를 지우며, 다른 한편으로는 위와 같은 사유가 발생한 경우 데이터제공자의 데이터수령자에 대한 책임 내지 정산에 관하여 정할 필요가 있다.
- 제공데이터가 요구되는 질을 갖추지 못한 경우의 책임범위를 규정할 필요가 있다. 데이터는 원료 내지 재료로 기능하므로 데이터의 흠이 상당한 파급효를 가질 수 있다. 데이터의 흠으로 데이터수령자가 공급한 상품이나 서비스에 흠이 생기거나 그로 인하여 고객 기타 제3자가 손해를 입어 데이터수령자에게 책임을 추궁하는 경우 그 책임을 인수할 의무를 지울 수 있다. 물론 이러한

의무는 원칙적으로 데이터제공자와 데이터수령자 사이에서만 효력이 있고 데이터수령자가 대외적으로 면책되는 것은 아니나 내부적 구상관계에 관한 법적 불확실성을 해소하는 데는 도움이 된다. 데이터수령자의 이용목적을 넘는 이용으로 인한 손해의 책임관계에 대하여도 명시하는 것이 바람직하다.

데이터거래계약의 구체적 내용(6): 제공데이터의 보안·비밀유지 등

- 제공데이터가 영업비밀인 경우 비밀관리성을 상실하면 데이터제공자가 임의의 제3자에 대하여 부정경쟁방지 및 영업비밀보호 등에 관한 법률에 따른 보호를 잃게 된다. 그러므로 데이터수령자에게 이 점을 주지시키고, 데이터수령자도 데이터제공자와 같은 수준의 비밀관리를 행하도록 데이터거래계약으로 강제할 필요가 있다. 이러한 관리에는 비밀분류 및 접근통제 등 여러 조치가 포함된다.
- 제공데이터가 개인정보인 경우 데이터수령자도 데이터제공자에 준하여 개인정보 보호법 등이 정하는 기술적·관리적 조치를 취하여야 하고, 비식별화·익명화된 정보인 경우 재식별위험을 주기적으로 평가하고 적절한 기술적·관리적 조치를 취하여야 하므로, 데이터제공계약에 이에 관한 사항을 포함시킬 필요가 있다. 데이터 유출에 대한 적시의 통지와 협력의무도 이에 포함된다.

기타 데이터제공관계의 형성

- 데이터거래를 위해서는 데이터거래계약 이외에도 데이터제공관계를 뒷받침하는 유형적·비유형적 인프라스트럭처가 필요하다.
- 데이터가 일회적으로 제공되는 것이 아니라 (실시간으로) 계속하여 제공되는 경우 데이터의 전송 등을 뒷받침하기 위한 회선과 접속경로 등의 확보 및 보안관리가 필요하다.
- 데이터수령자의 필요를 충족하면서도 불필요한 데이터공유가 일어나지 아니하도록 적절히 데이터를 가공할 수 있는 소프트웨어와 데이터제공자의 요청에 부응하는 접근통제·보안성 등을 갖춘 물리적 공간 및 하드웨어, 소프트웨어가 구축되어야 한다.
- 데이터거래는 많은 경우 데이터제공자와 데이터수령자 사이의 계속적 협력을 요구한다. 이는 데이터제공 자체를 위해서도 필요하지만 데이터와 관련된 위험을 감시·평가하고 대외적 대응이 필요한 경우 조치를 취하기 위해서도 필요하다. 미리 쌍방 접촉창구를 명확히 하고 데이터제공관계 전반을 파악하고 있는 담당자를 지정해놓을 필요가 있다.

4. 데이터결합 특유의 문제

데이터결합

- 데이터결합은 복수의 데이터셋을 결합하여 하나의 데이터를 만드는 것으로 각각의 데이터셋에서 동일한 사람, 장소, 시간 등을 식별하여 연결시킨다. 데이터결합을 통하여 각각의 데이터셋이 갖고 있는 정보가 합쳐져 의미 있는 정보를 창출할 수 있다.
- 결합의 대상이 되는 각 데이터가 비개인정보인 경우에는 특별한 문제가 없다. 그러나 각 데이터가 개인정보인 경우에는 결합의 각 단계에 문제가 생긴다.

개인정보인 데이터의 결합

- 개인정보인 데이터를 결합하여 비식별화·익명화된 비개인정보인 결합데이터를 생성하는 경우 서로 다른 데이터셋의 결합을 위하여 가명키를 활용하는 것이 보통이다. 즉 각각의 데이터보유자가 복호화가 불가능한 동일한 알고리즘을 활용하여 가명키로 대체하는 작업을 거친 뒤 이를 신뢰할 수 있는 제3자에게 제공하고 제3자가 데이터결합을 수행한 다음 가명키를 삭제한다.
- 각각의 데이터보유자가 복호화가 불가능한 동일한 알고리즘을 활용하여 가명키 대체하는 것만으로 개인정보가 비개인정보가 되는 것은 아니다. 복호화가 불가능하게 가명키로 대체하여도 원본 데이터를 갖고 있는 데이터보유자로서는 각각의 가명키가 어떤 정보주체에게 대응하는지 식별할 수 있는 경우가 많은데, 이때 각각의 데이터보유자가 신뢰할 수 있는 제3자에게 가명키로 대체된 데이터를 제공하는 것은 개인정보의 제3자 제공이 될 가능성이 있다. 가명키 대체로 비식별화·익명화되었다고 할 수 있는지에 대하여는 관련 분야 전문가의 조언을 구할 필요가 있다. 데이터결합기법 중에는 일대일 대응을 전제하지 아니하여 가명키 없이 결합하는 기법도 있는데, 이러한 기법을 활용하는 경우에는 개인정보 보호법 위반을 피할 수 있다. 그밖에 통계목적·학술목적 등을 위하여 가명키를 사용하여 개인정보를 제공하고 결합하는 것은 예외적으로 허용될 가능성이 있으나, 해석상 쟁점이 많고 판단도 미묘한 부분이 있으므로 관련 분야 전문가의 조언을 구함이 바람직하다.
- 데이터결합의 적법성 확보와 위험 관리에는 데이터보유자와 결합을 수행하는 제3자 사이, 그리고 복수의 데이터보유자 사이 제공되어야 하는 정보와 제공되어서는 안 되는 정보의 구별 및 정보제공 흐름의 통제가 관건적이다. 삼자간 데이터제공계약에서 이러한 측면을 명확히 규정할 필요가 있다.

결합데이터의 거래

- 결합데이터의 거래는 매매 유사의 관계일수도, 도급 유사의 관계일 수도 있다. 어느 경우든 데이터수령자에 대한 관계에서 데이터제공자, 즉 계약 당사자가

누구인가 하는 점과 대가를 어떻게 분배할 것인가 하는 점은 구분된다.

- 전자(前者)에 관하여는 결합 전 데이터제공자 중 일방 또는 쌍방이 결합데이터제공자가 되는 경우, 결합 전 데이터제공자 중 일방 또는 쌍방과 결합을 수행한 제3자가 공동으로 결합데이터제공자가 되는 경우, 결합을 수행한 제3자가 결합데이터제공자가 되는 경우가 있을 수 있다.
- 후자(後者)에 관하여는 결합을 수행한 제3자가 결합 전 데이터제공자들에게 일정한 대가를 지급하고 데이터를 제공받은 뒤 결합데이터거래로 인한 이익(과 손실)을 모두 부담하는 경우, 결합 전 데이터제공자들 중 일방 또는 쌍방이 결합데이터거래로 인한 이익(과 손실)을 모두 부담하는 대신 다른 결합 전 데이터제공자 및 결합을 수행한 제3자에게 일정한 대가를 지급하는 경우, 세 당사자가 조합 내지는 내적조합의 관계에서 손익을 분담하는 경우가 있을 수 있다.
- 어느 경우이든 대외적 거래와 책임의 주체 및 내부적 손익과 책임의 분담에 대하여 계약으로 명확히 규정할 필요가 있다.

5. 거래 데이터의 관리

데이터수령자에 대한 관계에서

- 데이터수령자가 제공받은 데이터를 계약에 따라 관리하고 있는지 주기적으로 또는 단기간인 경우 거래 종료 전 확인하여야 한다. 특히 개인정보인 경우에는 그 이외에 재식별위험의 변동 여부를 주기적으로 재검토할 필요가 있고, 관리적 조치와 관련하여 데이터수령자가 수행하여야 할 조치, 가령 접근통제, 접속기록유지, 주기적 교육 등이 실제 이루어지고 있는지 여부도 확인하여야 한다.
- 데이터수령자가 생성하여 축적한 파생데이터에 어떤 것이 있는지 및 데이터를 이용목적 내에서 이용하고 있는지 여부에 대하여도 주기적 감독이 필요하다.
- 데이터거래가 기간만료 또는 계약해지로 종료되는 경우 데이터가 적절히 삭제되었는지를 확인하여야 한다.

일반 대중에 대한 관계에서

- 개인정보인 데이터나 비식별화·익명화된 데이터의 경우 제공이 적법하다 하더라도 제공 사실에 대한 투명성이 대중의 신뢰 확보에 중요하다는 점을 인식하여야 한다. 제공 여부 및 목적, 제공으로 달성되는 가치, 상업적 이익 유무와 그 범위, 개인정보의 보호를 위한 조치 등에 대하여 적절히 공개하고 설명하는 것을 고려할 필요가 있다. 데이터수령자가 데이터를 유출하는 등 침해사고가 발생하는 경우 즉시 그 사실을 통지하고 데이터제공자의 대응도 수시 통지하여야 한다.
- 비개인정보인 데이터로 공익적 목적으로 데이터를 제공하는 경우에도 데이터공유정책의 가치에 대하여 일반 대중에게 충분히 이해시키는 것이 바람직하다.

서울대학교
인공지능정책
이니셔티브 안내

서울대학교 인공지능정책 이니셔티브는 인공지능과 관련된 다양한 사회경제적, 법적, 정책적 이슈들을 연구하고 논의하기 위해 시작된 서울대학교 법과경제연구센터의 프로그램입니다. ‘소셜랩(Social Lab)’ 개념을 지향하여, 여러 배경과 관심을 가진 분들 사이의 협업과 지속적인 대화를 추구합니다. 서울대학교 법학전문대학원의 교수와 임용 교수가 함께 이끌고 있습니다.

1. 발간물 안내

서울대학교 인공지능정책 이니셔티브의 주요 발간물은 이슈페이퍼와 워킹페이퍼가 있고, 비정기적으로 발간되는 단행본 및 학술행사 자료집 등이 있습니다. 이슈페이퍼와 워킹페이퍼 등의 자료들은 홈페이지를 통해 다운로드 받으실 수 있습니다.

2. 행사 안내

서울대학교 인공지능정책 이니셔티브의 주요 행사는 이슈페이퍼를 발표하고 논의하는 행사(상반기 및 하반기 각 1회) 그리고 국내외 연구자들을 초빙하여 진행하는 대규모 국제학술대회(연 1회) 등이 있습니다. 그 이외에 비정기적으로 진행하는 행사들도 있습니다.

3. 이슈페이퍼 2020-1

이번 이슈페이퍼는 서울대학교 인공지능정책 이니셔티브의 세 번째 이슈페이퍼로, 2020. 05. 07.에 열린 웨비나 행사를 위해 준비되었습니다.