

일반논문

빅데이터 규제의 새로운 패러다임 모색을 위한 연구^{†*}

- 개인정보 보호법제의 한계와 리스크규제체제로의
전환가능성에 대한 검토를 중심으로 -

A Search for New Regulatory Paradigm for Big Data:
Is a shift into the risk-based regulation feasible?

윤 혜 선(Hye-Sun Yoon)**

목차

- I. 서 론
- II. 빅데이터 고찰: 규제 대상의 관점에서
- III. 빅데이터에 대한 규제적 접근 방식 변화의 필요성 검토:
개인정보 보호법제에 의한 규제의 한계를 중심으로
- IV. 빅데이터에 대한 새로운 규제체제의 도입가능성 모색
- V. 결 론

<국문초록>

2013년부터 큰 관심을 받아온 빅데이터의 활용 수준이 아직까지 기대에 미치지 못하고 있다. 다양한 원인이 존재하는데, 빅데이터 활용의 활성화를 위한 제도적 정비가 미흡하다는 지적은 어김없이 그 목록에 등장한다. 실제 빅데이터의 활용을 둘러싼 개념적, 사회적, 법적, 윤리적 문제는 늘 제기되어 왔다. 빅데이터의 정의와 성질이 불분명하다는 지적에서부터, 빅데이터의 활용에 내재된 부작용으로서 사생활 침해, 부당한 차별, 위축효과, 필터버블의 우려, 약탈적 마케팅, 데이터에 대한 과도한 의존성 등 다양한 문

제가 제기되고 있는데, 이에 대한 명쾌한 규범적 기준이 제시되고 있지 못한 실정이다. 이 중에서 가장 주요하고 핵심적인 법적 과제는 개인정보 및 사생활 보호와 빅데이터 활용 간의 조화를 모색하는 문제이다. 이러한 이유로 빅데이터에 관한 국내 법적 논의는 개인정보 보호법제의 테두리를 벗어나고 있지 못한 실정이다.

이러한 현실에서 가장 이상적인 상황, 곧 빅데이터라는 신기술의 이용 활성화를 통해 다양한 편익과 혜택을 향유하고 기술 혁신을 촉진하면서 그에 따른 부작용을 가능한 최소화하고 그에 대한 책임 소재를 명확히 할 수 있는 규범적 환경을 조성하기 위해서는 발상의 전환이 필요한 것은 아닐까. 바로 이러한 문제의식 하에 본고는 빅데이터의 활용에 수반되는 부작용을 사회적으로 큰 편익을 제공하는 고도의 과학기술의 활용에 수반되는 부작용, 즉 리스크의 문제로 해석하여 빅데이터에 대한 규제적 대응을 리스크규제이론에 기초하여 새롭게 정립할 수 있는지 그 필요성과 가능성을 검토하였다.

본고의 구성은 다음과 같다. 먼저 빅데이터에 대하여 새로운 규제적 접근이 필요한지 여부에 대한 기초 조사로서 규제 대상으로서의 빅데이터를 고찰하였다(II). 빅데이터가 제공하는 편익과 부작용을 중심으로 빅데이터의 활용을 위한 법적 과제와 법적 대응방안에 대한 논의는 활성화되어 있는 편이지만, 현상으로서, 규제 대상으로서 빅데이터를 규범적 관점에서 검토한 연구는 부족한 편이기 때문이다. 이에 최근 해외에서 발표된 관련 연구결과를 기초로 빅데이터

† 투고일자 2018. 5. 10, 심사일자 2018. 5. 25, 게재확정일자 2018. 5. 28.

* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2015S1A5A8011167).

** 한양대학교 법학전문대학원 부교수 (Associate Professor, School of Law, Hanyang University).

의 정의, 독자성, 현황, 적용범위 등을 비롯하여 빅데이터가 제공할 수 있는 편익과 부작용 등을 검토하고 이것이 규제, 특히 새로운 규제 설계에 대하여 가지는 함의를 살펴보았다. 다음으로 빅데이터 규제에 대한 접근 방식의 변화의 필요성을 빅데이터의 특성과 현행 개인정보 보호법제의 체계와 내용을 비교하여 구체적으로 검토하였다(III). 마지막으로 빅데이터에 대한 새로운 규제체계의 도입가능성을 검토하였다(IV). 이를 위하여 먼저 빅데이터에 대한 새로운 규제적 접근을 시도함에 있어서 고려하여야 하는 쟁점들을 살펴보고, 리스크규제체제로의 빅데이터 규제 패러다임의 전환가능성을 검토한 후, 결론을 맺었다(V).

I. 서 론

바야흐로 우리는 데이터 시대에 살고 있다. 공행정, 정치, 경제, 금융 등 국가의 활동 부문뿐만 아니라 시민의 소소한 일상이 쉴 새 없이 수집·처리되는 데이터에 의해 변화되어 가고 있다. 상상이나 공상과학영화 속에서 그려왔던 모든 사물과 모든 사람이 기계와 네트워크로 연결되어 실시간 데이터를 수집·교환하는 세상의 도래가 가시화되었고, 웨어러블 기기, 인공지능 스피커, 자율주행자동차 등 서로 소통하는 기기들이 일상을 파고들기 시작하였다. 이미 휴대전화, 컴퓨터, 태블릿과 인터넷, 포털, 앱 등 다양한 기기, 네트워크, 플랫폼에서는 실시간 방대한 양의 데이터가 생성되고 있다. 이렇게 생성·수집·저장된 데이터는 스마트 기술이라고 불리는 알고리즘에 의해 조합·분석되고, 이렇게 분석된 데이터는 일상의 행위 패턴과 미래를 예측하고, 각종 절차를 간소화하며, 의사결정을 지원하고, 나아가 행동을 통제하는 데에 활용된다. 이처럼 금, 자본, 혈(blood), 가상 DNA, 지문 등으로 비유되는 데이터는 4차 산업혁명으로 지칭되는 첨단 기술 혁신의 요체이자 전제이다.

방대한 양의 데이터 또는 대량의 데이터를 수집·저장·조합·분석하는 기술로 이해되는 빅데이터는 이른바 데이터(주도)¹⁾ 혁신의 중심에서 현대 사회의 새로운 성장 동력으로 주목받고 있다. “21

세기의 원유”라고 은유되는 빅데이터의 활용 범위는 앞서 언급한 공행정, 정치, 경제, 금융 부문을 넘어 각종 산업부문, 사회, 교육, 보건의료, 과학기술, 문화예술, 스포츠, 사법(司法)영역에 이르기까지 다양할 뿐만 아니라 지속적으로 확장되고 있어서, 그 영향력과 활용에 대한 욕구가 더욱 커질 것으로 예상된다. 2013년 빅데이터에 주목하게 하는 하나의 계기를 만든 그의 저서에서 빅토르 마이어 쇠베르거 옥스포드 교수는 빅데이터를 지식정보혁명을 이끈 주역으로 인정받고 있는 구텐베르크의 활자에 비유한 바 있다.²⁾ 그의 주장에 따르면, 빅데이터는 인쇄술이 처음 나왔을 때와 같이 인류의 사고방식 자체를 바꿀 것이라고 한다. 빅데이터가 제공하는 편익과 효용은 아직 불확실하지만, 구글의 독감 예측 사례나 캘리포니아, 뉴욕 등 미국 몇몇 주(州)의 범죄예방 사례에서 보는 바와 같이 인터넷, 사물인터넷, 클라우드컴퓨팅, 인공지능 알고리즘 등 빅데이터가 다른 기술과 결합하여 시너지를 내게 되면 많은 긍정적인 변화가 생길 것으로 기대된다. 이에 OECD의 주요국을 중심으로, 공공 부문과 민간 부문 모두에서 빅데이터에 대한 투자와 시범사업이 늘어나고 있는 추세이다.³⁾ 우리나라의 경우에도 정부에서 빅데이터 산업을 육성하고 활용의 촉진시키

1) 데이터 주도 혁신(data-driven innovation)은 OECD에서 발전시킨 개념으로, 2010년 개시한 신성장동력: 지식기반자본 사업(New Sources of Growth: Knowledge-Based Capital) 제2기의 결과물에 의해 도출되었다. OECD, *Data-driven Innovation for Growth and Well-being, Interim Synthesis Report*, October 2014, <<https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf>> (최종방문일, 2018.4.1.) 참조.

2) 빅토르 마이어 쇠베르거/케네세 쿠기어/ 이지연 옮김, 『빅데이터가 만드는 세상』, 21세기북스, 2013 참조

3) European Data Protection Supervisor, *Report of workshop on Privacy, Consumers, Competition and Big Data 2 June, 2014*, 11 July 2014, <https://edps.europa.eu/sites/edp/files/publication/14-07-11_edps_report_workshop_big_data_en.pdf> (최종방문일 2017. 4. 1.); International Data Corporation(IDC), *Worldwide Semiannual Big Data and Analytics Spending Guide*, 2017, IDC는 빅데이터 및 비즈니스 분석 시장의 규모가 2016년 130.1억 달러에서 2020년에는 203억 달러 이상으로 증가할 것으로 예상하고 있다.; Bart van der Sloot/Sascha van Schendel, *Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study*, 7 (2016) JIPITEC 110 para 1.

기 위하여 다양한 정책적 시도를 하고 있는 것은 주지의 사실이다.⁴⁾

한편, 국내외에서 실시된 조사결과에 따르면, 2013년부터 본격적으로 주목받아온 빅데이터의 활용 수준은 아직까지 기대에 미치지 못하고 있는 상황이다.⁵⁾ 다양한 원인이 존재하는데, 빅데이터 활용의 활성화를 위한 제도적 준비가 미흡하다는 지적은 어김없이 그 목록에 등장한다. 실제 빅데이터의 활용을 둘러싼 개념적, 사회적, 법적, 윤리적 문제는 늘 제기되어 왔다. 빅데이터의 정의와 성질이 불분명하다는 지적에서부터, 빅데이터의 활용에 내재된 부작용으로서 사생활 침해, 부당한 차별, 위축효과, 필터버블의 우려, 약탈적 마케팅, 데이터에 대한 과도한 의존성 등 다양한 문제가 제기되고 있는데, 이에 대한 명확한 규범적 기준이 제시되고 있지 못한 실정이다. 이 중에서 가장 주요하고 핵심적인 법적 과제는 개인정보 및 사생활 보호와 빅데이터 활용 간의 조화를 모색하는 문제이다. 이러한 이유로 빅데이터에 관한 법적 논의는 줄곧 개인정보 보호법제의 구심력에 영향을 받고 있다.⁶⁾ 빅데이터의 활용을 통해 가치를 창출하고 싶은 측에서는 현행 ‘개인정보 보호법제’의 엄격성, 부적합성, 비효율성을 비판하고, 빅데이터 활용의 부작용을 우려하는 측에서는 현행 ‘개인정보 보호법제’의 불충분성, 실효성 부재, 집행력 미흡 등을 비판하고 있다. 법이론적 관점에서, 빅데이터의 적극적인 활용을 원하는 측에서는 ‘개인정보 보호법제’에 대

하여 개인정보의 재산적 가치와 효용에 기반한 재산권적, 시장논리적 접근을 취하고, 빅데이터의 활용에 대하여 다소 소극적인 또는 신중한 태도를 취하는 측에서는 개인정보의 인격적 가치에 기반한 인권적 접근을 정당하게 펼치고 있어서 양측 간에 접점을 찾는 것이 쉽지 않다.

이러한 현실에서 가장 이상적인 상황, 곧 빅데이터라는 신기술의 이용 활성화를 통해 다양한 편익과 혜택을 향유하고 기술 혁신을 촉진하면서 그에 따른 부작용을 가능한 최소화하고 그에 대한 책임 소재를 명확히 할 수 있는 규범적 환경을 조성하기 위해서는 발상의 전환이 필요한 것은 아닐까. 이를테면 현행 개인정보 보호법제의 패러다임과 다른 또는 이를 보완할 수 있는 규범적 패러다임을 모색하는 하는 방안을 생각해볼 수 있겠다. 그 자체로 신기술이면서 지속적으로 진화하는 빅데이터에 대한 법적 문제를 다른 시각에서 기술적, 사회심리적, 정책적, 경제적 환경에 맞게 재해석하고, 그에 적절한 새로운 규제 패러다임과 범리를 개발하는 시도는 고르디아스의 매듭(Gordian Knot)을 끊는⁷⁾ 효과가 있을 수 있다. 바로 이러한 문제의식 하에 본 연구는 빅데이터의 활용에 수반되는 부작용을 사회적으로 큰 편익을 제공하는 고도의 과학기술의 활용에 수반되는 부작용, 즉 리스크의 문제로 해석하여 빅데이터에 대한 규제적 대응을 리스크규제이론에 기초하여 새롭게 정립할 수 있는지 그 필요성과 가능성을 검토하는 것을 주요 목적으로 한다.

본고의 구성은 다음과 같다. 먼저 빅데이터에 대하여 새로운 규제적 접근이 필요한지 여부에 대한 기초 조사로서 규제 대상으로서의 빅데이터를 고찰한다(II). 빅데이터가 제공하는 편익과 부작용을 중심으로 빅데이터의 활용을 위한 법적 과제와 법적 대응방안에 대한 논의는 활성화되어 있는 편이지만, 현상으로서, 규제 대상으로서 빅데이터를 규범적 관점에서 검토한 연구는 부족한 편이다. 이에 최근 해외에서 발표된 관련 연구결과를 기초로 빅

4) 김진영, “빅데이터 분야 규제와 개선방안 검토 - 행정부의 정책방향과 입법과제를 중심으로”, 『IT와 법연구』, 제15집, 2017. 8. 157-190면; 최재경, “빅데이터 분석의 국내외 활용 현황과 시사점”, KISTEP InI 제14호, 한국과학기술기획평가원, (2016년 6월), pp.33-43; 한국정보화진흥원, “미래부-NIA, 2016년 빅데이터 시범사업 착수”, 보도자료, 2016. 6. 17. <<https://kbig.kr/portal/kbig/knowledge/news/press.page>> (최종방문일 2017. 4. 1.) 참조.

5) 최재경, 위의 글; 미래창조과학부/한국정보화진흥원/K-ICT 빅데이터센터, 2015년 BIG DATA 시장현황 조사, 2016; Bart van der Sloot/Sascha van Schendel, *Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study*, 7 (2016) JIPITEC 110 para 1.

6) 최근에는 저작물에 이용에 대한 동의, 일시적 복제 등 저작권법에 관한 쟁점도 부상하고 있다.

7) ‘고르디아스의 매듭을 끊는다’는 은유는 ‘복잡하고 난해한 문제를 지혜롭고 대담한 행동으로 푼다’는 의미로 사용되고 있다.

데이터의 정의, 독자성, 현황, 적용범위 등을 비롯하여 빅데이터가 제공할 수 있는 편익과 부작용 등을 검토하고 이것이 규제, 특히 새로운 규제 설계에 대하여 가지는 함의를 살펴본다. 다음으로 빅데이터 규제에 대한 접근 방식의 변화의 필요성을 빅데이터의 특성과 현행 개인정보 보호법제의 체계와 내용을 비교하여 구체적으로 검토한다(III). 마지막으로 빅데이터에 대한 새로운 규제체계의 도입가능성을 검토한다(IV). 이를 위하여 먼저 빅데이터에 대한 새로운 규제적 접근을 시도함에 있어서 고려하여야 하는 쟁점들을 살펴보고, 리스크규제체제로의 빅데이터 규제 패러다임의 전환가능성을 검토한 후, 글을 맺고자 한다(V).

II. 빅데이터 고찰: 규제 대상의 관점에서

지난 몇 년간 빅데이터에 집중된 관심에 비하여 규제 대상으로서 빅데이터의 개념이나 특성 등을 상세히 조사한 연구는 그리 많지 않은 편이다. 빅데이터에 대한 규제의 필요성과 목적을 정확히 인식하고, 현행 법적 수단의 적절성을 평가하고, 필요시 보다 적합한 규제 수단을 설계하기 위해서는 빅데이터의 정의, 실체, 활용 범위, 빅데이터가 제공하는 기회와 위험, 현행 법제도의 적정성 등에 대한 검토가 선행되어야 함은 물론이다. 이와 관련하여 최근 네덜란드에서 빅데이터 규제를 위한 기초 연구로서 전 세계 11개 국가를 대상으로 빅데이터 정책과 법제에 관한 비교법적 실증연구결과가 발표된 바 있다.⁸⁾ 이 연구보고서(이하 ‘네덜란드 연구

보고서’라 한다)에는 빅데이터 규제에 관하여 제기한 10가지 근본적인 질문에 대한 11개국 정부 관계자의 답변이 종합적으로 분석·정리되어 있다. 자료의 신뢰성 확보 등을 이유로 ‘정부’ 및 ‘정부 규제’의 관점에서 연구가 수행되었다는 한계가 있기는 하지만, 이 연구는 위에서 빅데이터 규제를 위하여 선행적 검토가 필요하다고 지적한 사항들을 직접적으로 비교적 상세하고 다루고 있을 뿐만 아니라 우리가 비교법적 연구 대상으로 삼는 미국, 독일, 프랑스, 영국, 일본, 호주 등 주요 기술선진국의 빅데이터 규제 현황을 포함하고 있다는 점에서 검토의 실익이 있다. 이하에서는 네덜란드 연구보고서의 관련 내용과 국내 개인정보 보호법제, 연구 문헌 등에 기초하여 규제 대상으로서의 빅데이터를 개념, 현상의 실제성, 기술적 독자성, 활용 범위, 활용 편익과 부작용을 중심으로 고찰하고, 각 사항에 대한 검토 내용이 규제에 대하여 가지는 함의를 검토하고자 한다.⁹⁾

1. 빅데이터의 정의

(1) 빅데이터의 개념 검토

빅데이터를 규제 연구의 시작은 그 개념에 대한 정의에서 출발한다. 대개의 신홍기술(emerging technologies)이 그러하듯 빅데이터에 대한 합의된 정의는 존재하지 않는다. 국가 혹은 단체마다 빅데이터의 주요 표지에 터잡아 조작적 정의를 사용하고 있는 것으로 보인다. 국제적으로 통용되고 있는 빅데이터의 정의는 크게 세 가지로 조사되었다.¹⁰⁾ 첫 번째는 유럽연합 제29차 작업반¹¹⁾의 정의이다.

8) Bart van der Sloot/Sascha van Schendel, *International and Comparative Legal Study on Big Data*, The Netherlands Scientific Council for Government Policy (WRR) Working Paper 20, 28-04-2016, <<https://www.wrr.nl/publicaties/working-papers/2016/04/28/international-and-comparative-legal-study-on-big-data>> (최종방문일 2018. 4. 1., 이하 ‘네덜란드 연구보고서’라 한다.) 이 연구의 대상이 된 11개 국가는 호주, 브라질, 중국, 프랑스, 독일, 인도, 이스라엘, 일본, 남아프리카 공화국, 영국, 미국 등이며, 네덜란드 Tilburg 대학 연구진은 이 연구를 위하여 각국의 빅데이터 관련 법제와 정책에 대한 문헌 연구와 빅데이터 업무를 담당하는 각국 정부 담당자에 대한 설문 조

사를 실시하였다.

9) 아울러 네덜란드 연구보고서의 연구진은 연구결과를 기초로 “빅데이터 규제설계를 위한 10가지 질문”이라는 제목의 논문도 발표하였다. Bart van der Sloot/Sascha van Schendel, 앞의 논문. 여기에서는 네덜란드 연구보고서와 해당 연구진의 위의 논문의 내용을 중심으로 빅데이터 규제에 관한 근본적이고 다양한 쟁점들을 검토하기로 한다.

10) 네덜란드 연구보고서, pp.15-17.

11) Article 29 Working Party로 약칭되며, The Working Party on the Protection of Individuals with regard to the Processing of Personal Data이 정식명칭이다. Directive 95/46/EC on the

이 정의에 따르면, 빅데이터는 “가용 정보와 정보의 자동화된 이용의 기하급수적 성장”이며, “기업, 정부, 기타 대규모 단체가 보유하고 있는 방대한 디지털 데이터의 집합으로서, 컴퓨터 알고리즘을 사용하여 광범위하게 분석되는 데이터의 집합”으로, “보다 일반적인 경향이나 상관관계를 파악하기 위하여 이용될 수 있는, 그러나 개인에게 직접 영향을 미치지 위하여 처리될 수 있는 것”으로 기술된다.¹²⁾ 두 번째 정의는 유럽개인정보보호감독관(EDPS)¹³⁾이 제안한 것으로서, 이에 따르면, 빅데이터는 “다수의 출처에서 고속으로 생성된 다양한 종류의 대용량 데이터로서, 그 처리와 분석에 있어서 새롭고 보다 강력한 프로세서와 알고리즘이 요구되는 것”으로 정의한다.¹⁴⁾ 세 번째 정의는, 가장 널리 알려진 것으로 2012년 가트너 보고서(Gartner Report)가 제안한 3V 모델 혹은 3V 이론이다. 이러한 별칭이 붙은 이유는 동보고서가 빅데이터를 정의할 때 세 가지 핵심표지, 곧 규모(데이터 양, volume), 속도(데이터 처리속도, velocity) 및 다양성(데이터 유형 및 출처의 범위, variety)의 증가에 초점을 맞추었기 때문이다.¹⁵⁾

국내의 경우, 비록 빅데이터에 관한 독자적인 법

protection of individuals with regard to the processing of personal data and on the free movement of such data 제29조에 의해 설치된 개인정보보호 및 프라이버시에 관한 유럽연합 차문기구이다. 유럽연합 제29조 작업반에 관한 상세한 내용은 <<https://epic.org/privacy/art29wp/>> (최종방문일 2018. 4. 1.) 참조.

- 12) Bart van der Sloot/Sascha van Schendel, 앞의 논문, pp.112-113.
- 13) EDPS로 약칭되며 European Data Protection Supervisor이 정식명칭이다. 유럽연합의 독립적인정보보호 감독기구이며, 기능, 권한 등에 관한 상세한 내용은 EDPS 홈페이지 <https://edps.europa.eu/about-edps_en> (최종방문일 2018. 4. 1.) 참조. EDPS는 빅데이터가 모두 개인정보는 아니지만, 디지털 경제에서 시장 참여자들의 대규모의 개인정보의 수집과 거래에 대한 의존도가 점차 커질 것이라고 예상하고 있다.
- 14) <https://secure.edps.europa.eu/EDPSWEB/edps/Consultation/big_data See also: https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Big%20data/14-07-11_EDPS_Report_Workshop_Big_data_EN.pdf> (최종방문일 2018. 4. 1.).
- 15) <<http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>> (최종방문일 2018. 4. 1.).

률은 존재하지 않지만, 현행 법령에서 빅데이터에 관한 정의 규정을 찾을 수 있으며,¹⁶⁾ 다음의 세 가지 정의가 존재한다. 첫째, “(초(超)) 대용량의 정형 또는 비정형의 데이터세트”¹⁷⁾, 둘째, “정보처리능력을 가진 장치를 통하여 공개정보 및 이용내역정보 등을 처리(수집, 생성, 저장, 조합, 분석, 그 밖에 이와 유사한 행위를 말한다)하여 생성되는 정보 및 이로부터 가치를 추출하고 결과를 분석하는 기술”¹⁸⁾, 셋째, “디지털환경에서 생성되는 정형 또는 비정형의 수치, 문자, 영상 등의 대량 데이터의 집합 및 이로부터 가치를 추출하고 결과를 분석하는 기술”¹⁹⁾. 이러한 빅데이터의 조작적 정의에 따르면, 우리나라에서는 다종의 대용량의 데이터 또는 특정한 정보를 처리하여 생성된 데이터 그 자체와 이러한 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 빅데이터의 주요 징표로 삼고 있음을 알 수 있다.

네덜란드 연구보고서는 11개국에서 사용되는 빅데이터의 정의를 모두 검토한 후 적어도 현 시점에

- 16) 관련하여 방송통신위원회는 빅데이터 활용 활성화를 위하여 2014년 말 「빅데이터 개인정보보호 가이드라인」을, 2016년 7월 1일 「개인정보 비식별 조치 가이드라인」을 발표한 바 있으나 이들 가이드라인에는 빅데이터에 대한 명시적인 정의규정이 존재하지 않는다. 더욱이 「개인정보 비식별 조치 가이드라인」은 조문 형식이 아니라 안내서 형태로 구성되어 있다. 다만, 「빅데이터 개인정보보호 가이드라인」 제1조(목적)가 “이 가이드라인은 공개된 개인정보 또는 이용내역정보 등을 전자적으로 설정된 체계에 의해 수집·저장·조합·분석 등 처리하여 새로운 정보를 생성함에 있어서 이용자의 프라이버시 등을 보호하고 안전한 이용환경을 조성하는 것을 목적으로 한다”고 규정하고 있어서 이 규정을 통해 적어도 방송통신위원회는 “(공개된 개인정보 또는 이용내역정보 등을 포함한) 정보를 전자적으로 설정된 체계에 의해 수집·저장·조합·분석 등 처리하여 새로운 정보를 생성”하는 것을 빅데이터의 개념으로 이해하고 있다는 추론을 할 수 있다.
- 17) 「행정절차법 시행령」 제25조의2 제2항 제2호, 「행정안전부와 그 소속기관 직제」 제12조 제3항 제21호, 「과학기술정보통신부와 그 소속기관 직제」 제13조 제3항 제22호 등.
- 18) 「부산광역시 빅데이터 활용 및 빅데이터산업 육성에 관한 조례」 제2조 제1호, 「충청북도 빅데이터 활용에 관한 조례」 제2조 제2호, 「충청남도 데이터기반행정 빅데이터 활용에 관한 조례」 제2조 제2호 등. 이 정의는 방송통신위원회의 「빅데이터 개인정보보호 가이드라인」에서 도출되는 빅데이터의 개념을 조작적 정의로 규정한 것이다. 앞의 주 16) 참조.
- 19) 경기도, 경상남도, 경상북도, 광주광역시, 부산광역시, 의왕시, 전라북도(아날로그 환경도 포함), 전주시, 제주특별자치도 및 창원시 「빅데이터 활용에 관한 조례」 제2조, 「대구광역시 빅데이터 활용에 관한 조례」 제2조 제1호 등.

서 빅데이터를 정확하게 정의하는 것은 불가능하다는 결론을 내렸다.²⁰⁾ 대신, 각국에서 사용하고 있는 빅데이터의 개념을 모두 분석한 후, 빅데이터를 정의할 때 데이터의 ‘수집’, ‘분석’, ‘이용’의 빅데이터 처리의 세 단계 단계를 고려할 필요가 있다고 권고하였다.²¹⁾ 먼저, 빅데이터는 ‘대용량(volume)’의 정형 및 비정형 데이터를 ‘다양한 출처(variety)’에서 ‘수집’하는 것을 핵심 징표로 한다. 또한 빅데이터의 개념에는 ‘분석’이 포함되는데, 분석의 구성 요소에는 알고리즘, 기계학습, 통계적 상관관계 등의 ‘분석기법’과 분석의 ‘속도(velocity)’가 포함된다. 빅데이터 분석의 결과는 대개 예측적인 성질(predictive analytics)을 가지며, 개인 수준에서가 아니라 일반적인 수준 또는 집단의 수준에서 설명되는 특징을 가진다. 나아가 이러한 분석결과는 주로 프로파일링을 통해 ‘이용’된다. 정리하면, 위에서 기술한 내용의 ‘수집’, ‘분석’, ‘이용’은 빅데이터 여부를 식별하는 핵심 표지라고 할 수 있다. 즉, 이들 중 어느 하나의 표지도 가지지 않은 현상은 명백히 빅데이터에 해당하지 않는 반면, 세 가지 표지를 모두 가지는 현상은 빅데이터에 해당한다. 네덜란드 연구보고서의 조사 결과에 따르면, 현재 사용되는 빅데이터의 정의들은 이 세 표지 중 하나 이상을 포함하고 있지만, 세 표지를 모두 포함하고 있는 정의는 아직 존재하지 않는다. 전술한 우리 빅데이터의 정의도 ‘수집’ 또는 ‘수집’ 및 ‘분석’의 두 표지를 포함하고 있다. 즉, 현재 우리나라를 포함하여 주요국에서 사용하고 있는 빅데이터의 정의는 이 세 표지가 구성하는 입체적 스펙트럼의 사이 어딘가에 위치하고 있으며, 이 위치는 시간이 지남에 따라 기술 발전 등 다양한 요인에 의하여 변할 수 있다.

(2) 규제적 함의

빅데이터를 정의하기 어려운 이유는 위에서 살

펴본 것처럼 단지 기술이 발전 중에 있으며(emerging), 빅데이터의 활용 범위가 확대되고 있어서(dynamic) 개념범위의 확장가능성이 존재하기 때문만은 아니다. 물론 이것이 주요한 이유이기도 하지만, 빅데이터로 분류되기 위해서 데이터의 규모가 얼마나 방대해야 하는지, 비정형 데이터가 반드시 포함되어야 하는지, 비정형 데이터는 어느 정도 포함되어야 하는지, 출처는 얼마나 다양해야 하는지 등에 대한 선명한 기준, 즉 빅데이터의 주요 표지 - 수집, 분석, 이용 - 를 구성하는 세부 요소들의 총족 여부를 판단하는 기준조차 존재하지 않기 때문이기도 하다.

빅데이터가 다양한 구성요소를 포함하는 표지들을 가진다는 사실과 빅데이터의 개념범위의 확장가능성, 빅데이터의 개념을 정의하는 것이 상술한 이유 등으로 쉽지 않다는 사실은 빅데이터 규제에 대하여 다음과 같은 점을 시사한다. 첫째, 규제를 위하여 빅데이터를 정확하고 정교하게 정의하는 것은 매우 도전적인 과제이다. 이는 특히 빅데이터에 대한 새로운 규제 체계를 모색하는 경우 그 체계의 설계뿐만 아니라 구축 속도, 도입 시점 등에 적지 않은 영향을 미칠 수 있다. 둘째, 규제를 위해 채택되는 빅데이터의 조작적 정의는 빅데이터의 기술, 효용 등의 확장가능성에 영향을 미칠 수 있다. 따라서 매우 신중한 접근이 필요하다. 셋째, 빅데이터는 수집, 분석, 이용 세 단계를 모두 포함하며, 이들은 상호 긴밀하게 연결되어 있으므로, 빅데이터 규제를 설계하거나 집행할 때에는 이러한 연관성을 고려하여야 한다.

2. 빅데이터의 실존성

(1) 빅데이터 현상의 실존성 검토

빅데이터 활용에 대한 사회의 지속적인 관심이 높은 상황에서 빅데이터의 실존성을 운운하는 것이 다소 뜬금없어 보일 수 있다. 빅데이터의 실존성이란 빅데이터의 활용 현황에 관한 문제 제기라고 할 수 있다. 이러한 문제 제기가 있는 이유는 빅데이

20) 네덜란드 연구보고서, p.17.

21) 위와 같음.

터에 쏟아진 세간의 관심이나 빅데이터 활성화를 위한 정부의 다양한 정책적 제도적 노력에 비해, 빅데이터의 활용실적이 미미하기 때문이다. 이러한 사정은 다른 국가들도 비슷한 것으로 보인다.²²⁾ 이에 대하여 네덜란드 연구보고서는 빅데이터를 실체가 없는 허구나 과장광고로 간주하는 태도를 지양하고, 지금으로부터 5년이나 10년 후 중요한 역할을 할 것이며 정부와 기업의 운영방식과 시민의 일상에 큰 영향을 미칠 것으로 기대되는 일종의 트렌드로 볼 필요가 있다고 권고하였다.²³⁾ 달리 말하면, 빅데이터의 영향력이나 효용은 지금이 아닌 가까운 미래에 보다 명확하게 파악할 수 있다는 의미이다.

(2) 규제적 함의

이상과 같은 빅데이터의 현황이 빅데이터 규제에 대하여 가지는 함의는 규제를 고려할 때 이러한 ‘트렌드’가 반영되어야 하며, 이러한 긍정적인 기대가 효과적으로 실현될 수 있도록 미래 지향적인 규제적 접근을 하여야 한다는 것이다. 특히, 빅데이터의 실존성에 대한 지적은 규제의 시기와 방식에 대하여 시사하는 바가 크다. 규제의 시기와 관련하여, 빅데이터의 효과가 명확해지는 시점에 비로소 정부가 개입을 한다면, 규제의 집행가능성을 보장하기 어려울 수 있다. 그 즈음에는 수년 전에 시작된 빅데이터 프로젝트들의 부정적인 결과들이 구체화될 것이고, 한참 진행된 프로젝트를 조정하거나 취소하는 데에 많은 비용과 어려움이 따를 것이기 때문이다. 규제의 집행가능성과 실효성을 높이기 위해서는 정부가 보다 신속하게 개입할 필요가 있는데, 이 경우 규제의 방식과 수단 선택에 있어서 신중한 접근이 요구된다. 법적 기준이 모호하거나 범위반의 우려가 존재하는 경우 빅데이터의 활용에 대한 유인이 감소하게 될 것임은 자명하므로, 빅데이터의 혁신과 활용을 장려할 수 있도록 명확하고 적정

한 수준과 내용의 규제를 마련하여야 할 것이다. 또한 규제의 수준이나 내용이 적절하지 못한 경우 기술 발전에 왜곡을 가져올 수 있으므로 적정한 규제수단을 마련하기 어렵다면 반응적이거나 유연한 규제적 접근을 고려할 필요가 있다.

3. 빅데이터의 독자성

(1) 빅데이터의 기술 또는 현상으로서의 독자성 검토

빅데이터에 대한 관심은 클라우드컴퓨팅, 사물인터넷, 인공지능 기술의 발전과 함께 증가해 온 것이 사실이다. 빅데이터는 일반적으로 공공데이터 정책(Open Data), 데이터 재활용(Re-Use), 사물인터넷, 스마트²⁴⁾, 프로파일링, 알고리즘, 클라우드컴퓨팅과 함께 언급되는 것으로 관찰된다.²⁵⁾ 국내에서도 첨단 정보통신기술 외에도 정부의 공공데이터의 개방 및 이용 활성화 정책과 데이터기반 행정정책²⁶⁾ 적극적으로 추진되면서 빅데이터의 이용 활성화에 대한 논의가 더욱 부각된 면도 있다. 이에 과연 빅데이터가 독립적인 현상 내지 기술인지 짚어 볼 필요가 있다. 국내에서는 사물인터넷, 클라우드컴퓨팅, 인공지능 등 개별 기술별로 규범 체계를 구축해 나가는 추세이고²⁷⁾ 공공데이터 정책의 경우 「공공데이터의 제공 및 이용 활성화에 관한 법률」이라는 규범체계가 이미 구축되어 있기 때문이다.

기술이 발전하면서 이와 같은 인위적인 구별과

22) 위와 같음, pp.20-22.

23) 이에 대한 근거로 미국, 영국, 프랑스, 남아프리카 공화국 및 독일 정부가 추진하고 있는 대규모 빅데이터 프로젝트의 사업의 주요 내용과 투자규모를 제시하였다. 위와 같음, pp.21-22.

24) 사물인터넷의 활성화와 상시 소통할 수 있는 기기들이 증가함에 따라 스마트 기기, 스마트 전력망, 스마트 시티 등과 같은 스마트 제품과 서비스도 늘어가는 추세이다. 이러한 연결고리 안에서 스마트도 빅데이터와 긴밀한 관계를 형성하고 있다.

25) Bart van der Sloot/Sascha van Schendel, 앞의 논문, pp.114-115 & 125.

26) 정부는 2017. 12. 28. 「데이터기반 행정 활성화에 관한 법률안(정부제출, 의안번호:2011077)」을 국회에 제출하였다. 이 법률안은 데이터기반의 과학적 행정체계를 구축하여 행정의 책임성과 국민 삶의 질을 제고하며, 공공기관 간 데이터의 공유를 활성화하고 데이터 통합 관리체계를 마련하는 것을 목적으로 한다(제1조).

27) 클라우드컴퓨팅의 경우, 클라우드컴퓨팅의 발전 및 이용을 촉진하고 클라우드컴퓨팅서비스를 안전하게 이용할 수 있는 환경을 조성하기 위하여 2015. 3. 27. 「클라우드컴퓨팅 발전 및 이용자 보호에 관한 법률」이 제정되었다.

매칭의 타당성이 흐려질 수 있으나 일반적으로 빅데이터는 데이터의 ‘수집’과 관련하여 사물인터넷과, 데이터 ‘분석’ 및 ‘이용’과 관련하여 알고리즘, 기계학습 등 인공지능과, ‘수집’, ‘분석’, ‘이용’ 과정 전반에서 클라우드컴퓨팅²⁸⁾과 긴밀하게 연결되어 있다. 사물인터넷이나 분석 알고리즘의 경우는 빅데이터의 개념적 기술적 구성요소로 볼 수도 있다.²⁹⁾ 이러한 관점에서 접근하면 빅데이터는 독자적인 현상이나 기술이라기보다는 현재 빠르게 발전하고 있는 다수의 데이터 기반 기술들을 포괄하는 현상 내지 상위의 기술이라고 할 수도 있다.

공공데이터 정책은 정부 및 공공기관의 데이터는 공유되어야 한다는 관념에 기초한다. 이 정책은 본래 공공부문에 대한 투명성을 높여서 시민과 언론에 대한 정부의 부당한 공권력 행사에 대한 통제를 강화하려는 목적으로 추진된 것이다.³⁰⁾ 최근 공공데이터 정책은 공공데이터의 이용과 재활용이라는 관념으로 변화·발전되고 있다. 「공공데이터의 제공 및 이용 활성화에 관한 법률」은 제정 목적을 “공공기관이 보유·관리하는 데이터의 제공 및 그 이용 활성화에 관한 사항을 규정함으로써 국민의 공공데이터에 대한 이용권을 보장하고, 공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함을 목적으로 한다”고 선언하고 있다.³¹⁾ 전통적인 공공데이터 정책과 공공데이터의 이용·재활용 정책 사이에 한 가지 중요한 차이가 존재하는데, 바로 후자의 정책은 민간 기업과 개인의 공공데이터의 상업적·영리적 활용을 촉진하는 것을 목적으로 한다는 점이다.³²⁾ 이러한 차이가 가

지는 중요한 함의는 공공데이터의 이용·재활용 정책에 따라 공공데이터는 본래 수집된 목적 이외의 목적으로 사용될 수 있으며, 그것이 이 정책의 핵심 목표라는 점이다. 우리나라 등 다수 국가에서 사용하는 빅데이터의 정의에는 데이터의 재활용과 분석을 통해 이차적 또는 새로운 가치를 추출하는 관념이 포함되어 있으며, 이것이 바로 빅데이터의 핵심 영업모델이자 존재 의의라고 할 수 있다.

(2) 규제적 함의

빅데이터는 다양한 개념, 현상, 기술들과 긴밀하게 연결되어 있고, 특히 정보통신기술과 정보처리 기술을 구성요소로 하는 개념이라고 할 수 있다. 따라서 빅데이터 규제를 적용·집행하거나 새롭게 설계하는 경우, 빅데이터가 별개의 독립적인 현상이 아니라 그 개념상 현재 발전 중에 있는 다수의 기술적, 정책적(법적), 사회적 변화와 매우 밀접한 관련이 있다는 점이 고려되어야 한다. 이러한 이유로 빅데이터 규제는 관련된 다양한 현상, 기술 및 제도와 이들의 성장속도, 동태성 등이 포함된 지형도 안에서 총체적인 접근방식을 추구하거나³³⁾ 이와 반대로 관련된 현상, 기술, 제도의 개별적인 규범체계 안에서 다루는 방식을 추구할 필요가 있다.

4. 빅데이터의 활용 범위

(1) 빅데이터의 활용 범위 검토

빅데이터의 활용 범위에는 사실상 제한이 없다.

28) 「클라우드컴퓨팅 발전 및 이용자 보호에 관한 법률」 제2조 제1호에 따르면, 클라우드컴퓨팅이란 집적·공유된 정보통신기기, 정보통신설비, 소프트웨어 등 정보통신자원(이하 “정보통신자원”이라 한다)을 이용자의 요구나 수요 변화에 따라 정보통신망을 통하여 신속적으로 이용할 수 있도록 하는 정보처리체계를 의미한다.
 29) 예를 들면, 영국에서는 빅데이터에 정보주체가 의식적으로 제공한 데이터뿐만 아니라 사물인터넷 기기 등으로부터 수집되거나, 다른 데이터에서 도출되거나, 분석과 프로파일링을 통해 유추된 개인정보도 포함된다고 정의하고 있다. 네덜란드 연구보고서, p. 18.
 30) Bart van der Sloot/Sascha van Schendel, 앞의 논문, p.114.
 31) 「공공데이터의 제공 및 이용 활성화에 관한 법률」 제1조.
 32) 「공공데이터의 제공 및 이용 활성화에 관한 법률」 제14조는 공공

데이터 이용 활성화를 촉진하기 위하여 정부가 추진할 수 있는 사업 및 활동에 대하여 규정하고 있는데, 특히 2016. 1. 6. 신설된 동조 제2항은 “정부는 공공데이터 이용 활성화를 촉진하기 위하여 공공데이터를 활용한 창업을 촉진하고 창업자의 성장·발전을 위하여 필요한 지원을 할 수 있다.”고 규정하고 있다. 나아가 중복·유사 서비스 개발·제공의 방지를 규율하는 제15조의2는 “공공기관의 장은 공공데이터를 활용하여 개인·기업 또는 단체 등이 제공하는 서비스와 중복되거나 유사한 서비스를 개발·제공하여서는 아니 되”며(제1항), “공공기관의 장은 중복·유사 서비스의 개발·제공을 방지하기 위한 방침을 마련하고, 새로운 서비스를 개발·제공하기 전에 중복·유사 투자 여부를 확인하여야 한다”(제2항)고 규정하고 있다.
 33) 네덜란드 연구보고서, p.20.

활용 기반만 잘 조성되어 있다면 빅데이터는 거의 모든 영역의 모든 업무에 적용될 수 있다. 그러나 규제의 관점에서는 빅데이터가 활용되는 목적, 분야·대상, 범위 등에 따라 국민의 권익에 미치는 영향이 상이하고, 이는 규제의 방식, 내용, 강도 등 규제 설계 전반에 영향을 미치기 때문에 적절하고 합리적인 규제를 위해서 빅데이터의 활용 분야, 목적 등에 대한 검토는 매우 중요하다.

빅데이터의 이용 현황을 기초로 활용 분야는 크게 세 가지로 범주화될 수 있다.³⁴⁾ 첫 번째 범주는 정부가 특정 업무를 위하여 빅데이터를 활용하는 경우이다. 국가정보기관의 정보서비스, 경찰의 경찰행정업무, 세무관청의 조세행정업무, 공공기관의 경제 정책 수립 등에 빅데이터를 활용하는 것이 이 범주에 속한다. 두 번째 범주는 민간 또는 공공기관·준정부기관이 경영 목표나 업무상 목표를 달성하기 위하여 빅데이터를 활용하는 경우이다. 이들은 대개 사업에 수반되는 리스크를 관리하고, 통계적 상관관계를 찾고, 맞춤형 서비스와 광고를 제공하기 위해 빅데이터를 활용한다. 대학이나 연구기관에서도 연구 관련 목적으로 빅데이터를 이용한다. 세 번째 범주는 정부와 민간 기업이 각각 시민과 고객에 대한 서비스를 개선하고 편의를 증진하기 위하여 빅데이터를 활용하는 경우이다. 정부와 기업 활동의 투명성을 높이고, 개인정보 처리에 대한 국민의 통제권을 강화하기 위하여 빅데이터를 활용하는 경우 등을 포함한다.

한 눈에도 각 범주별로, 그 범주 내에서 이루어지는 활동별로 빅데이터에 대하여 다른 규제적 접근이 필요하다는 사실을 알 수 있다. 예를 들면, 세 번째 범주의 경우에는 빅데이터의 활용이 시민의 이익에 부합하기 때문에 상대적으로 적은 또는 약한 규제적 접근이 가능해 보인다. 그러나 첫 번째 범주와 같이 행정기관에서 특정 업무 수행의 효율성과 실효성을 제고하기 위하여 빅데이터를 활용하는 경우에는 사정이 달라진다. 또한 첫 번째 범주

내에서도 어떠한 기관이 또는 어떠한 목적으로 빅데이터를 이용하느냐에 따라 규제적 고려가 상이해진다. 예컨대, 미래 예측, 경제 정책 개발, 역학조사 등과 같이 국민의 권리·의무관계에 직접적인 영향을 미치지 않는 업무에 빅데이터를 이용하는 것은 규제의 관점에서 상대적으로 문제의 소지가 적다고 할 수 있다. 이러한 업무에는 주로 일반적인 패턴과 통계적 상관관계 분석이 활용되기 때문이다. 그러나 구체적인 사실을 바탕으로 특정인을 조사하여 그에 대하여 적절한 조치를 취하여야 하는 경찰이 대량의 데이터를 처리하여 ‘일반적인 패턴을 도출하는 것을 목적’으로 하는 빅데이터를 활용하는 경우는 사정이 다르다. 일반적인 패턴분석을 통해 도출된 일반적인 프로파일을 특정 개인에게 적용할 경우 불일치의 위험과 그에 따른 권리침해의 가능성이 존재하기 때문이다. 민간 기업이 빅데이터를 이용하는 경우에는 특히 다음과 같은 두 가지 사항이 공통적으로 지적되고 있다. 하나는 인종, 건강상태, 종교 등과 같은 민감정보가 부당하게 이용될 수 있다는 점과 다른 하나는 빅데이터를 개인에 대하여 이용할 경우 민감정보의 이용여부와 관계없이 그 결과가 치명적일 수 있다는 사실이다. 빅데이터 정책 마련을 위한 각종 보고서나 언론에서 자주 언급된 대표적인 사례처럼 금융기관이나 보험회사가 각각 대출 대상자와 대출 조건을, 보험가입 대상자와 가입조건을 정하기 위하여 빅데이터를 이용하여 개발한 리스크 프로파일을 자동 적용하는 경우 개인과 사회에 상당한 부정적인 결과를 초래할 수 있다.

(2) 규제적 함의

빅데이터 이용 기반이 잘 조성되어 있다면, 사실상 그 활용 범위에 한계가 없으며, 실제로도 활용 범위가 지속적으로 확대되고 있다는 현실을 감안할 때, 빅데이터 규제를 설계하거나 집행하는 경우 다음과 같은 사항을 고려할 필요가 있다. 첫째, 빅데이터를 이용하는 주체와 목적, 그러한 빅데이터의 이용이 개인의 권익에 미치는 잠재적인 영향, 빅데이터에 이용되는 데이터의 유형과 분석기법, 일반

34) 위와 같음, p.23.

적인 프로파일과 특정 개인 간의 불일치 가능성 등을 고려하여야 한다. 빅데이터가 경찰, 국가정보원, 군당국 등에서 개인을 대상으로 한 특정 임무의 수행을 위해 이용되는 경우, 일반 행정기관에서 경제, 사회를 위한 일반 정책 개발을 위해 이용되는 경우에 비해 개인에게 더 큰 영향을 미칠 수 있음은 물론이다.³⁵⁾ 이에 둘째, 빅데이터를 사용하는 조직의 유형과 빅데이터를 사용하는 구체적인 목적들을 구별해내는 작업이 필요하다. 이렇게 구분된 조직의 유형과 목적별로 상이한 규제적 접근을 모색하여야 할 것이다.

5. 빅데이터 활용에 따른 편익

(1) 빅데이터 활용의 편익 검토

지금까지 알려진 빅데이터의 유용성은 대용량의 데이터 분석을 통해 기존에 알지 못했던 지식이나 가치, 해안을 도출하거나 패턴을 추출하여 현황 분석이나 이상현상 감지, 미래예측 등에 활용할 수 있다는 점에서 찾을 수 있다.³⁶⁾ 이를 통해 의사결정의 합리성과 정책의 실효성을 제고하고, 각종 업무처리절차를 간소화하여 업무 능률과 시스템운영의 효율성을 높이며, 맞춤형 서비스의 제공가능성을 확대하고, 변화에 대한 사전대응능력을 질적으로 향상시킬 수 있을 것으로 기대된다. 그러나 적어도 현재 빅데이터가 구체적으로 어떠한 편익을 가져올지 명확한 그림이 현출되어 있지는 않다.

빅데이터의 품질과 성능, 활용 목적, 활용 분야 등에 따라 편익과 기회的内容이 다를 것이다. 다만, 한 가지 분명한 사실은 빅데이터의 활용이 민간부문에서 특히 활성화될 것이고, 창업 촉진, 기업 이익 제고 등을 통해 경제적 가치와 일자리를 창출

하여 상당한 경제 성장을 가져올 것이라는 것에 주요국의 모든 이해관계자가 견해를 같이 하고 있다는 것이다.³⁷⁾

(2) 규제적 함의

규제의 관점에서 빅데이터 활용에 따른 편익은 규제의 정당성과 내용, 수준을 결정짓는 핵심적인 요소로 작용한다. 특히, 빅데이터에 대한 규제에 대하여 비례원칙, 이익형량 등을 적용하여 그 적법성과 합목적성, 적정성을 심사하는 경우 빅데이터 활용의 구체적인 편익의 가치는 더욱 중요하게 된다. 따라서 빅데이터 규제 설계 시, 빅데이터 활용의 구체적인 혜택과 실익에 대한 면밀한 검토가 필요하다.

6. 빅데이터 활용에 따른 부작용

(1) 빅데이터 활용의 부작용 검토

빅데이터 활용이 가져올 수 있는 부작용으로 개인정보 및 사생활 보호 권리 침해의 우려, 데이터 결정론, 부당한 차별가능성, 힘의 불균형, 위축효과, 필터버블 등이 지적되고 있다.³⁸⁾

1) 개인정보 및 사생활 보호 권리 침해 우려

개인정보 및 사생활 보호 권리의 침해의 우려는 빅데이터의 활용과 관련하여 이미 주지된 가장 중요한 위험이며, 이 권리들의 보호를 위한 개인정보 보호법제의 주요 원칙들이 빅데이터와 상충된다는 점에서 정당하게 제기되는 위험이다. 개인정보 및 사생활 보호권 침해 우려의 문제는 이들 권리의 인권적, 헌법상 기본권적 성질뿐만 아니라 그 침해의 결과가 빅데이터 활용에 따른 다른 부작용들에 대하여 직간접적인 원인적 관련성을 가진다는 점에서 매우 중요하고 신중하게 다루어져야 한다. 무엇보다 현행 「개인정보 보호법」, 「정보통신망 이용촉진

35) 이러한 이유로 다수의 국가에서 경찰의 수사 업무 등에 빅데이터를 활용하는 방안에 대해서 회의적인 것으로 조사되었다. 네덜란드 연구보고서, p.24 참조.

36) 최경진/정준현/구태연/지성우/김도승/성준호, 『빅데이터 환경에서 개인정보보호 강화를 위한 법·제도적 대책 방안 연구』, 개인정보 보호위원회, 2012, 14-23면; 고유흠, “빅데이터와 개인정보보호”, 『지식재산동향』, Vol 21, 한국지식재산연구원, 2014, 12, 70-83면, 73-75면 등 참조.

37) 네덜란드 연구보고서, p.27.

38) 최경진 외, 앞의 연구보고서, pp. 29-32; 네덜란드 연구보고서, pp.28-34.

및 정보보호 등의 관한 법률(이하 「정보통신망법」이라 한다.)에서 개인정보 및 사생활을 보호하기 위하여 기본 원칙으로 규정한 최소수집의 원칙, 목적외 활용금지 원칙, 데이터품질 보장 원칙, 투명성 원칙 등³⁹⁾이 빅데이터와 공립할 수 없는 긴장관계를 형성하기 때문에 이 권리들의 침해의 우려가 현실화될 개연성이 상당히 높은 편이라고 할 수 있으므로 시급한 검토가 요청되는 문제이기도 하다.⁴⁰⁾

2) 데이터결정론과 부당한 차별 우려

빅데이터의 활용을 활성화하려는 정책적 의지의 근원에는 더 많은 데이터를 수집·분석하면 보다 더 정확하고 합리적인 의사결정을 할 수 있을 것이라는 기대가 있다. 그러나 “더 많은 데이터 = 더 많은 지식 = 더 좋은 결정”이라는 등식이 반드시 성립하는 것은 아닐 수 있다. 오히려 많은 데이터로 인해 결정과정에서 더 많은 혼란과 거짓 양성(false positives)이 발생하여 결정 장애나 불합리한 결정을 초래할 수도 있다. 또한 빅데이터는 인과관계가 아닌 상관관계에 기초한다. 이로 인해 특히 개인에 대한 빅데이터의 이용이 부당한 결과를 가져올 수도 있다. 통계적 상관관계는 확률적 개념이므로, 어떠한 현상이 발생할 개연성이 높다고 해서 그 현상이 100% 발생할 것이라는 결론을 보장할 수 없다. 빅데이터에 대한 의존도가 점점 더 커지면, 향후 우리의 자유의지에 따른 실제 행위가 아닌 데이터가 말해주는 개연성이 높은 예상 행위에 기초하여 의사결정이 이루어지는 이른바 데이터의 독재(Dictatorship of Data)가 초래될 수도 있다.⁴¹⁾

빅데이터에 대하여 가장 많이 지적되는 부작용 중 하나는 자동처리되는 결정이나 예측 분석이 개인에게 부정적인 영향을 미칠 수 있다는 것이다.⁴²⁾ 분석 알고리즘은 중립적이지 않다. 알고리즘은 특정한 목적을 달성하기 위한 데이터, 연결, 추론, 해

석, 포함여부에 관한 기준 등에 대한 선택적 결과물이기 때문이다. 따라서 알고리즘을 통한 분석이 핵심인 빅데이터는 기존의 사회적 편견이나 고정관념을 통합하여 사회적 배제와 계층화, 부당한 차별을 강화할 위험도 내포하고 있다.

3) 힘의 불균형 심화 우려

현대 사회에서 기업, 특히 대기업의 행위에 대하여 개인이 통제력이나 영향력을 가지기는 매우 어렵다. 그런데 빅데이터 활용이 지금보다 활성화되면 특히 기업과 소비자 간의 협상력과 힘의 불균형이 더욱 커질 수 있다는 우려가 제기된다. 기업은 소비자의 개인정보를 수집하고, 이를 기초로 여러 가지 가치와 패턴을 도출해 내는데 점점 더 역량을 집중할 것이다. 이에, 예컨대, 취업, 은행대출, 보험 등의 거래가 소비자에게 불리한 방식으로 이루어질 우려가 있으며, 약탈적 마케팅, 소비자 유형에 따른 차별적 가격결정이 이루어질 가능성도 제기된다.

4) 위축효과 우려

기업이 개인이 온·오프라인의 일상에서 생성하는 데이터를 수집·분석하여 맞춤형 상품이나 서비스를 제공한다면, 그러한 행위에 대하여는 당장 사생활 침해, 부당한 차별 등의 이슈가 제기될 것이다. 그러나 이러한 상황이 장기간 지속되면, 가령 10년 후에는, 우리가 알게 모르게 제공한 다양한 정보를 기초로 예컨대 보험상품 등의 가입이 건강 조건 등을 이유로 거절당할 수도 있다. 이러한 상상적 혹은 실제 실현이 임박한 결과는 현재의 시민의 일상생활과 사회참여에 있어서 부정적인 위축효과를 가져올 수 있다. 인터넷, 통신망, 플랫폼에 남기는 흔적과 방문기록 등이 향후 구직 등 개인의 삶에 중요한 다양한 결정들에 영향을 미칠 수 있다는 우려와 식별가능성, 프로파일링 등에 대한 우려 때문에 온라인에서 다른 지식을 검색·제공하거나 다양한 견해를 피력하는 것을 주저하게 될 수 있기 때문이다. 국가의 빅데이터 이용과 관련하여서도, 정보 수집의 출처나 사용용도에 대한 불확실성은

39) 「개인정보 보호법」 제3조, 「정보통신망법」 제23조 제2항, 제30조 등.

40) 관련하여 자세한 내용은 다음 항에서 다루기로 한다.

41) 네덜란드 연구보고서, p.33.

42) 위와 같음.

정부에 대한 국민의 신뢰를 위협하고, 사회적 위축 효과를 초래하여, 결과적으로 개방적이고 건전한 민주주의의 근간을 훼손시킬 우려가 있다. 요컨대, 공·사를 불문하고 빅데이터의 이용에 있어서 사생활 보호가 미흡하게 이루어지게 되면 시민의 개인적, 사회적, 경제적, 문화적, 정치적 활동에서부터 표현의 자유, 사회 참여에 이르기까지 위축효과를 야기하여 궁극적으로 건전한 자유민주주의 사회 유지·발전에 부정적인 영향을 미칠 수 있다.

5) 필터버블 효과 우려

포털, 앱 등 플랫폼이나 웹사이트의 개인화 현상과 개인의 인터넷, 모바일이용 행위에 기반한 맞춤형 미디어콘텐츠와 뉴스서비스 제공 행위는 다양한 견해와 아이디어를 교환하는 공론의 장의 형성과 유지 조건에 영향을 미칠 수 있다. 이 문제도 건전한 민주주의의 전제에 대한 것으로, 이러한 현상을 이른바 반향실 내지 메아리효과 또는 필터버블이라고 한다.⁴³⁾ 필터버블은 개인화된 알고리즘에 의해 형성되는 개인적인 정보공간을 말하는 것으로, 그 내용은 정보주체의 성향, 선호에 따라 검색 등 알고리즘 필터에 의하여 다르게 채워진다. 필터버블 안에서는 시민이 자신의 관점과 가치를 확인해 주는 콘텐츠에만 노출되므로, 자신의 의견과 가치만이 증폭되게 된다. 이러한 필터버블에 갇히게 되면 제한된 정보를 제공받고, 다른 관점과 견해를 가진 다양한 사람들에 대한 노출이 제한되며, 개인의 생각과 관점을 소통하고 교환할 수 있는 범위도 제한되므로 자기계발의 기회가 한정되고, 그에 따라 민주주의가 훼손될 수 있다는 우려가 지적되고 있다.

(2) 규제적 함의

이상에서 보는 바와 같이 빅데이터 활용에 따른 부작용들의 대부분은 사생활 보호가 미흡하여 발생하거나 부당한 차별의 문제로 수렴되는 것으로 보인다. 이러한 부작용은, 반복하지만, 빅데이터 이용

주체, 분야, 목적에 따라 개인과 사회에 미치는 영향이나 그 발현의 내용과 정도가 상이할 수 있다. 예를 들면, 빅데이터 이용에 따른 부당한 차별은 항상 문제가 되지만, 경찰의 공권력행사에 있어서 개인에 대한 부당한 차별행위는 기업의 맞춤형 광고행위에 의한 소비자간 차별에 비해 더욱 중대하고 심각하다고 할 수 있다. 따라서 빅데이터 규제의 관점에서 특히, 사생활 침해와 부당한 차별의 위험은 중요하고 정당하게 평가·고려되어야 한다. 그러나 이러한 부작용을 빅데이터 이용에 따른 일반적이고 보편적인 문제로 인식하고 일률적으로 접근하려는 태도는 지양되어야 한다. 그것보다는 개별 분야별로 세분화하여 개별적으로 접근하고, 빅데이터 활용이 제공하는 구체적인 편익과의 합리적인 이익형량 등을 통해 규제의 구체적 타당성을 확보하려는 자세가 필요하다.

7. 소결

이상으로 규제 대상으로서 빅데이터에 관한 기초 사항들을 검토해 보았다. 빅데이터의 정의와 개념징표, 빅데이터의 이용현황 및 활용가능성, 사물인터넷, 클라우드컴퓨팅, 인공지능 등 첨단 정보통신·처리기술 및 공공데이터 정책과 같은 데이터 이용 활성화 정책과 빅데이터 간의 긴밀한 연관성, 빅데이터 활용에 따른 편익과 부작용 등 빅데이터의 다양한 특성들은 현행 개인정보 보호법제 하에서 개인정보 및 사생활 보호와 빅데이터 이용 활성화 간의 조화를 모색하는데 본질적인 한계가 존재한다는 사실을 직간접적으로 확인시켜 주었다. 뿐만 아니라 새로운 규제체계를 강구함에 있어서도 빅데이터의 이러한 특성들이 상당한 도전을 제기할 것이라는 것도 확인할 수 있었다. 이제 항을 바꾸어 빅데이터의 규제에 대한 접근 방식의 변화의 필요성을 현행 개인정보 보호법제의 체계와 내용을 가지고 보다 구체적으로 검토한다.

43) 위와 같음, p.34.

Ⅲ. 빅데이터에 대한 규제적 접근 방식 변화의 필요성 검토: 개인정보 보호법제에 의한 규제의 한계를 중심으로

빅데이터는 「개인정보 보호법」, 「정보통신망법」 등으로 대표되는 개인정보 보호법제의 적용을 받는다.⁴⁴⁾ 개인정보 보호법제는 현행 법질서 하에서 빅데이터의 이용 활성화와 개인정보 및 사생활 보호 간의 조화를 도모할 수 있는 거의 유일한 법적 수단이다. 그럼에도 불구하고 개인정보 보호체계의 근간을 이루는 주요 기본 원칙들과 빅데이터는 본질적으로 긴장관계에 형성한다. 이러한 사실은 두 가지 함의를 가지는데 하나는 개인정보 보호법제에도 불구하고 빅데이터의 활용이 사생활 및 개인정보 보호에 위협을 가할 수 있다는 점이고, 다른 하나는 현행 개인정보 보호 규제체계가 빅데이터에 대하여 적합하지 않을 수 있다는 것이다. 이러한 논의의 연장선에서 현행 개인정보 보호법제는 빅데이터의 활용에 의해 발생하는 손해나 권리침해에 대한 적절한 권리구제 수단을 제공하지 못한다는 지적도 제기된다. 이에 법적 규제가 아닌 다른 방식의 규제적 접근을 시도하지는 제안도 일각에서 이루어지고 있다. 이하에서는 빅데이터와 개인정보 보호법제가 채택한 기본원칙들 간의 긴장관계, 개인정보 보호법제 하에서 권리구제의 한계, 법적 규제 방식의 한계 등을 차례로 검토한다.

1. 개인정보 보호법제의 기본 원칙과 빅데이터의 관계 검토

(1) 최소수집의 원칙과 빅데이터의 관계

개인정보처리자는 개인정보의 처리 목적을 명확하게 하여야 하고, 그 목적에 필요한 범위에서 최소한의 개인정보만을 적법하고 정당하게 수집하여

야 한다(최소수집의 원칙).⁴⁵⁾ 최소수집의 원칙은 가능한 한 적은 양의 데이터를 수집하여야 하며, 데이터의 양이 수집 목적과 관련하여 과도하지 않아야 한다는 것을 의미한다. 나아가 수집된 목적이 달성된 경우 해당 개인정보는 지체 없이 파기되어야 한다(개인정보 삭제의무).⁴⁶⁾

최소수집의 원칙은 빅데이터와 정면으로 상충된다. 빅데이터의 핵심 개념은 최대한 많은 데이터를 수집하고, 이미 수집된 데이터로부터 언제나 새로운 가치를 찾을 수 있다는 것이기 때문이다. 수집된 목적이 달성된 경우 개인정보를 삭제하여야 하는 의무도 수집된 데이터에는 항상 새로운 목적이 부여될 수 있다는 빅데이터의 관념과 충돌한다. 요컨대 최대한 많은 정보 수집이 요체인 빅데이터는 필연적으로 최소수집의 원칙과 긴장관계를 형성할 수밖에 없으며, 개인정보 및 사생활 보호 권리 침해 우려의 발생은 이러한 관계의 논리필연적인 결과라고 할 수 있다.

(2) 목적외 활용금지의 원칙과 빅데이터의 관계

개인정보처리자는 개인정보의 처리 목적에 필요한 범위에서 적합하게 개인정보를 처리하여야 하며, 그 목적 외의 용도로 활용하여서는 아니 된다(목적외 활용금지의 원칙 또는 목적제한원칙).⁴⁷⁾ 본래 수집된 목적 외의 다른 목적을 위한 개인정보의 추가 처리를 금지하는 목적외 활용금지의 원칙은 개인정보의 2차 사용을 원칙적으로 허용하지 않는다. 문제는 빅데이터를 위한 데이터 수집 단계에서는 대개 고정된 목적이 존재하지 않는다는 사실이다. 방대한 양의 데이터가 단순히 수집을 위해 수집될 뿐이며, 그렇게 수집된 데이터의 가치나 잠재적 용도는 수집된 이후에 비로소 분명해진다. 더욱이 빅데이터 분석과정에서는 여러 종류의 데이터 베이스가 다른 종류의 데이터(집합)와 결합되거나

44) 「개인정보 보호법」, 「정보통신망법」 외에도 개인정보 보호를 규율하는 다양한 실정법들 - 「위치정보의 보호 및 이용 등에 관한 법률」, 「신용정보의 이용 및 보호에 관한 법률」, 「생명윤리 및 안전에 관한 법률」 등 - 이 존재하기는 하지만 본고의 논의의 목적상이 두 법률을 중심으로 검토하기로 한다.

45) 「개인정보 보호법」 제3조 제1항 및 제16조 제1항, 「정보통신망법」 제23조 제2항 등.

46) 「개인정보 보호법」 제21조 제1항, 「정보통신망법」 제29조 등.

47) 「개인정보 보호법」 제3조 제2항 및 제21조 제1항, 「정보통신망법」 제24조 및 제29조

병합되는 것이 다반사인데, 이 과정에서 데이터가 수집된 본래 목적이 상실된다. 빅데이터는 데이터의 양도 물론 중요하지만 데이터의 재활용과 분석을 통해 새로운 가치를 도출하는 것이 무엇보다 중요하다. 이러한 이유로 목적외 활용금지의 원칙은 빅데이터와 개인정보 보호 사이의 긴장관계를 야기하는 원인으로 가장 많이 언급되고 있다.⁴⁸⁾

(3) 데이터품질 보장의 원칙과 빅데이터의 관계

개인정보처리자는 개인정보의 처리 목적에 필요한 범위에서 개인정보의 정확성, 완전성 및 최신성이 보장되도록 하여야 한다(데이터품질 보장의 원칙).⁴⁹⁾ 이는 개인에 대하여 생성된 프로파일과 그에 기초한 결정의 합리성과 정확성을 보장하기 위한 것이다. 그런데 빅데이터는 주로 개인보다는 집단의 수준에서, 회고적인 분석이 아니라 미래에 대한 예측적 확률분석이 주로 이루어지기 때문에 데이터의 정확성이나 최신성에 대한 일정한 오류 여유(margin of error)를 허용한다.⁵⁰⁾ 뿐만 아니라 분석 대상 데이터의 대부분이 일반적인 수준에서 정확한 그림을 제공할 수 있다면 정보처리자에게 있어서 높은 품질의 데이터 확보의 문제는 그리 중요하지 않다. 더욱이 정보처리자들은 “질 보다는 양”이라는 표현처럼, 낮은 품질의 데이터(‘dirty data’)를 가지고 작업하는 것에 점점 더 익숙해지기 때문에⁵¹⁾ 빅데이터에 있어서 데이터 품질 보장의 원칙의 엄격한 적용은 과도한 법적 요구가 될 수 있다.

(4) 투명성의 원칙과 빅데이터의 관계

개인정보처리자는 개인정보 처리방침 등 개인정보의 처리에 관한 사항을 공개하여야 하며, 열람청구권 등 정보주체의 권리를 보장하여야 한다(투명성의 원칙).⁵²⁾ 투명성의 원칙은 개인정보 보호법제

에서 개인정보열람권⁵³⁾, 개인정보의 정정·삭제권⁵⁴⁾, 정보주체 이외로부터 수집한 개인정보의 수집 출처 등 고지 의무⁵⁵⁾ 등으로 구체화되어 있는데⁵⁶⁾ 이들 역시 빅데이터와 일정한 긴장관계를 형성한다.

먼저, 빅데이터 환경에서 정보주체는 자신의 개인정보가 수집된다는 사실 자체를 인지하지 못하는 경우가 많기 때문에 개인정보열람권과 그에 기초한 개인정보의 정정·삭제권 등을 행사하지 못할 가능성이 크다. 또한 투명성의 원칙은 개인정보처리자에 대해서도 수집출처 등 고지 의무에 대한 이행가능성의 문제를 제기한다. 개인정보처리자도 특히 빅데이터 처리과정에서 서로 다른 데이터베이스가 결합되거나 데이터의 재활용이 있는 경우 개인정보의 주체나 수집 출처를 모를 수 있고, 정보주체의 연락처를 알 수 없는 경우도 있을 수 있기 때문이다. 그럼에도 불구하고 빅데이터 처리를 통해 개인에 대한 프로파일이 생성될 수 있으며, 이에 기초한 의사결정이 개인에게 적잖이 심각한 영향을 미치는 불투명한 자동화된 방식으로 이루어질 수 있다는 문제가 상존한다.

빅데이터가 공공데이터의 제공 및 이용 활성화 정책과 연계되어 있는 경우에도 투명성의 원칙과 상충될 가능성이 존재한다. 이 정책에 따라 다양한 기관, 특히 민간 사업자들에 의해 공공데이터가 빅데이터의 방식으로 처리되는 경우, 적절한 투명성 보장조치 없거나 그러한 조치가 존재한다고 하더라도 빅데이터의 특성상 그러한 조치를 현실적으로 준수하기 어렵기 때문이다. 경우에 따라서는 법적 근거 없이 다양한 출처에서 수집한 데이터의 집합과 데이터베이스들이 결합될 수도 있다.⁵⁷⁾

(5) 데이터 안전관리의 원칙과 빅데이터의 관계

개인정보처리자는 개인정보의 처리 방법 및 중

48) 네덜란드 연구보고서, pp.28-29.

49) 「개인정보 보호법」 제3조 제3항.

50) 네덜란드 연구보고서, p.30.

51) 위와 같음.

52) 개인정보 보호법, 제3조 제5항, 「정보통신망법」 제27조의2 및 제30조 제2항 등.

53) 「개인정보 보호법」 제35조.

54) 「개인정보 보호법」 제36조.

55) 「개인정보 보호법」 제20조.

56) 「정보통신망법」 제30조, 제30조의2 등.

57) 네덜란드 연구보고서, p.31.

류 등에 따라 정보주체의 권리가 침해받을 가능성과 그 위험 정도를 고려하여 개인정보를 안전하게 관리하여야 하며,⁵⁸⁾ 개인정보의 분실·도난·누출·변조 또는 훼손을 방지하기 위하여 법령에서 정한 기술적·관리적 조치를 하여야 한다(데이터 안전관리의 원칙).⁵⁹⁾ 개인정보 보호에 있어서 데이터 안전관리의 원칙 또는 기술적·관리적 조치 원칙은 매우 중요하다. 빅데이터 환경에서 이 원칙은 앞의 네 원칙과 다른 의미의 긴장관계를 형성한다. ‘대용량’의 개인정보를 처리하는 빅데이터에 대하여 데이터 보안을 강화하여야 한다는 요구와 함께 엄격하고 효과적인 기술적·관리적 보안조치에 대한 요구는 항상 당연히 따르는 것이기 때문이다.

2. 권리구제의 한계

현행 개인정보 보호법체계는 개인정보에 대한 개인의 권리를 강조하고 있고, 이를 강화하는 추세이다. 개인정보 보호의 기본법이라 할 수 있는 「개인정보 보호법」은 동법의 목적이 개인정보의 처리 및 보호에 관한 사항을 정하여 “개인의 자유와 권리를 보호하고, 나아가 개인의 존엄과 가치를 구현”하는 것에 있다고 천명하고 있다(제1조). 2011년 3월 29일 최초 제정되어 거의 매년 개정되어 온 동법은 지속적으로 개인정보에 대한 개인의 권리를 강화해 오고 있다. 우리 개인정보 보호법제의 진화에 많은 영향을 끼치고 있는⁶⁰⁾ 유럽연합의 개인정보 보호법제는 2018년 5월 25일을 기점으로 일반개인정보보호규정(General Data Protection Regulation, 이하 ‘GDPR’이라 한다)의 체제로 전환되었는데, 새로운 GDPR에는 삭제권(잊혀질 권리)(제17조),

개인정보 이동권(제20조), 처리반대권(제21조), 설명요구권(전문 71, 제13조(2)(f), 제14조(2)(g)) 등이 개인정보에 대한 개인의 권리를 강화하기 위한 수단이 다수 포함되어 있다.

개인의 권리는 의무를 수반한다. 이는 개인의 권리를 보호하고, 권리의 침해가 발생한 경우 권리를 행사할 수 있도록 보장하기 위함이다. 오히려 개인에게 맡기는 ‘개인적인 권리와 의무’를 중심으로 한 개인정보 및 사생활 보호의 규범체계와 권리구제 체계가 앞에서 설명한 특성을 가지는 빅데이터의 시대에도 유지될 수 있을지 의문이다. 빅데이터가 활성화된 환경에서는 개인이 손해나 권익의 침해를 입증, 아니 인식하는 것이 현재보다 더 어려워질 것으로 예상된다. 무과실책임체제가 도입된다고 하더라도, 개인이 자신의 권리가 침해되고 있다는 사실을 인지조차 하지 못하는 경우도 있을 수 있고, 자신에 관한 데이터가 자신도 모르게 수집되었다는 사실 자체를 모를 수 있다. 빅데이터를 위한 데이터의 수집은 실시간 방대하게 이루어져서 개인이 개별 데이터 처리절차에서 자신의 개인정보가 포함되어 있는지 여부를 확인하는 것이 사실상 불가능하므로 개인정보 처리가 적법하게 이루어졌는지 여부를 확인하는 것도 현실적으로 어렵다. 종합해 보건대, 빅데이터 환경에서 ‘개인’의 법적 권리에 기초한 법원 등을 통한 권리구제는 권리구제의 내용과 범위, 책임의 주체 등에 관한 논의를 차지한다 하더라도 청구 자체가 쉽지 않을 것으로 예상된다. 즉, 빠른 정보통신·처리기술의 발전으로 현행 법질서 하에서 빅데이터에 의한 개인정보 및 사생활 침해에 대한 우려가 커지고 있는데 그에 대한 실효적인 권리구제의 가능성은 그와 반대로 기술발전의 속도와 같은 속도로 감소하고 있는 이러한 현실은, 빅데이터의 이용 활성화와 개인정보 및 사생활 보호 간의 조화를 도모하기 위한 다양한 시도를 저지 또는 지연시키는 요소로 기능할 가능성이 매우 농후하다.

58) 「개인정보 보호법」 제3조 제4항.

59) 「정보통신망법」 제28조.

60) 일례로 4차산업혁명위원회가 2차 및 3차 규제·제도혁신 해커톤을 통해 합의·도출된 ‘개인정보의 보호와 활용의 조화’ 방안들은 유럽연합 일반개인정보보호법(EU GDPR)을 주로 참조한 것이다. 안선희, “4차산업위 ‘가명정보, 상업적 목적 통계에 이용 가능’”, 한겨레신문, 2018. 4. 5. 자 기사<<http://www.hani.co.kr/arti/economy/it/839264.html#csidx1833090997c4ac4b6b65b4457e4c5a>> (최종방문일, 2018. 4. 28.).

3. 법적 규제의 방식의 한계

전술한 바와 같이 현행 개인정보 규제체제는 권리와 의무를 중심으로 한 법적 체제이다. 빅데이터는 이러한 규제의 ‘방식의 적절성’에 대하여도 의문을 제기한다. 그 이유는 첫째, 데이터 처리의 범위가 국경을 초월하고 있으며, 앞으로 국가간 데이터 이전은 더욱 확대될 것이기 때문이다. 이러한 이유로 국가간 협약, 국제적 협약의 필요성이 제기되고 있으나⁶¹⁾ 국제 협약의 체결 그 자체, 나아가 그러한 협약의 법적 구속력과 집행력을 확보하는 것이 현실적으로 매우 어렵다는 것은 익히 경험을 통해 체득한 바이다. 둘째, 규제체제의 문제이다.⁶²⁾ 빅데이터 기술이 빠르게 변화하고 있기 때문에, 특정한 법규정의 적절성과 합리성이 유지되기 어렵고, 그로 인해 규제 준수가 회피될 가능성이 높아지거나 예측하지 못한 문제가 발생할 가능성이 커진다. 신흥기술 규제에서 공통적으로 나타나는 이러한 규제지체 현상은 빅데이터의 경우에도 피할 수 없다. 이에 대처하는 효과적인 방안은 반응적이고 유연한 규제 접근을 하는 것이다. 셋째, 빅데이터 활용에 따른 부작용의 사회이슈화 현상이다. 빅데이터 처리와 관련하여 제기되는 부작용의 문제들 중 다수가 사회적 이슈로 발전·확대되는 경향이 있어서 개인 차원의 권익 보호를 중심으로 발전된 특정 법리를 가지고 이 문제들을 다루는 데에 한계가 있다는 지적이다.⁶³⁾ 이러한 이유로 국내외적으로 법적 규제를 대체하거나 보완하는 대안적인 규제방식에 대한 논의가 활발하다. 자율규제, 실무수칙, 윤리적 지침의 개발 등이 주로 논의되고 있는데, 이러한 방식들도 익숙한 한계들이 존재하므로 여전히 법적 구속력이 있는 규제에 대한 필요성이 존재한다.

61) 네덜란드 연구보고서, p.32.

62) 규제지체에 대한 논의는 윤혜선, “‘신흥기술 규제(emerging technologies regulation)’ 연구방법론 개발을 위한 기초연구 - 기술규제 시 고려요소를 중심으로 -”, 『행정법연구』 제49호, 2017. 6., 103-128면 참조.

63) 네덜란드 연구보고서, p.32.

IV. 빅데이터에 대한 새로운 규제체제의 도입가능성 모색

1. 빅데이터에 대한 새로운 규제체제의 도입을 둘러싼 법적 쟁점 검토

(1) 필요성

빅데이터를 위한 새로운 규제체제의 도입이 필요할까? 빅데이터의 특성을 고려할 때 현행 개인정보 보호법제가 개인정보 보호에 있어서도 빅데이터의 활용에 있어서도 불충분하다는 점은 인정된다. 특히, 빅데이터는 개인정보 보호를 위한 기본 원칙들과 일정한 긴장관계를 형성하는데, 이러한 이유로, 경우에 따라서는, 개인의 삶과 공동체에 야기될 수 있는 빅데이터의 부작용에 대한 적절한 사전 관리가 어려울 수 있으며, 실현된 손해에 대한 권리구제 역시도 쉽지 않을 수 있다는 점은 앞에서 살펴본 바와 같다. 또한 현행 개인정보 보호법제가 특히 민간 부문의 빅데이터 이용 활성화에 상당한 부담으로 작용하고 있다는 것도 공공연한 사실이다. 보다 근본적인 문제는 많은 이해관계자들이 빅데이터에 대하여 현행 법체계가 실제로 어떻게 적용되고 해석될 것인지 또는 되어야 하는지에 대해 확신하지 못하고 있다는 점이다. 이와 같은 법적 불명확성과 불확실성은 적어도 두 가지 극단적인 부작용을 초래할 수 있다.⁶⁴⁾ 한편으로는 이해당사자들이 법위반에 대한 우려 때문에 사실상 합법적인 기술 혁신의 산물과 데이터의 사용을 포기할 수 있다는 것이고, 다른 한편으로는 이해당사자들이 현행 규제체제에 부합하지 않는 특정 기술을 사용하기 위하여 이러한 회색상태를 이용 혹은 악용할 수 있다는 것이다. 새로운 규제체제의 도입을 고려한다면, 그 체제가 빅데이터에 관한 위와 같은 법적 쟁점들에 대해 해결책을 제공할 수 있는지 여부와 그 방식의 적절성에 대해 신중하게 평가할 필요가 있다.

64) Bart van der Sloot/Sascha van Schendel, 앞의 논문, p.127.

(2) 새로운 빅데이터 규제의 출발점: 개인정보?

빅데이터를 위한 새로운 규제체제의 구축은 다음의 두 가지 이유로 인해 특히 어려울 것이라는 주장이 제기된다.⁶⁵⁾ 첫 번째 이유는 빅데이터에 대한 새로운 규제적 접근에 있어서 좋은 출발점을 선택하는 것이 어렵기 때문이고, 다른 이유는 빅데이터 처리에 있어서 개인정보처리자를 특정하는 것이 어렵기 때문이다. 먼저, 빅데이터 규제의 출발점에 관한 사항을 검토하고, 두 번째 이유는 목차를 바꾸어 살펴본다.

현행 개인정보 보호법제에 의해 확립된 규제체제는 개인과 개인의 권익을 기초로 정보주체와 개인정보처리자의 법률관계를 규율하고 있다. 그 바탕에는 개인정보 보호, 사생활보호와 같은 인권과 기본권이 자리 잡고 있는데 이들은 ‘개인정보’, 즉 살아있는 개인을 알아볼 수 있는 정보(해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것을 포함한다)⁶⁶⁾라는 개념을 그 출발점으로 한다. 그러나 앞에서 살펴본 바와 같이, 빅데이터 처리는 일반적으로 개인 수준에서의 데이터의 수집·분석·이용보다는 대량의 데이터의 집적, 일반 패턴 찾기, 그룹 프로파일 등 그룹 수준에서의 데이터 처리에 주로 활용된다. 그러므로 빅데이터에 대한 새로운 규제체제를 개인과 개인정보를 출발점으로 하여 설계할 것인지에 대한 진지한 검토가 필요하다. 통계적 상관관계와 일반적인 패턴, 그룹 프로파일 등은 그 자체로 개인정보에 해당하지 않을 수 있지만, 국민 개개인의 생활에 상당한 영향을 미치는 각종 의사결정에 사용될 수 있다. 따라서 최근 메타데이터의 이용이 활성화되고 있는 추세이다.⁶⁷⁾

65) 위의 논문, pp.123-124.

66) 「개인정보 보호법」 제2조 제1호. 「정보통신망법」에서는 ‘개인정보’를 “생존하는 개인에 관한 정보로서 성명·주민등록번호 등에 의하여 특정한 개인을 알아볼 수 있는 부호·문자·음성·음향 및 영상 등의 정보(해당 정보만으로는 특정 개인을 알아볼 수 없어도 다른 정보와 쉽게 결합하여 알아볼 수 있는 경우에는 그 정보를 포함한다)”로 정의하고 있다(제2조 제1항 제6호).

67) Bart van der Sloot/Sascha van Schendel, 앞의 논문, pp.123-124. 다만, 메타데이터가 언제 개인정보에 해당하는지

그러나 빅데이터 처리과정에서 개인정보나 개인 프로파일이 개인정보가 아닌 다른 데이터의 이용·조합·분석을 통해 생성될 수도 있다는 사실도 간과해서는 아니 된다. 디지털 경제에서 개인정보는 핵심 자원이므로 개인정보의 방대한 수집과 거래에 대한 경제 주체들의 의존도가 높아지고 있는 추세이다. 또한 빅데이터 처리에 항상 개인정보가 포함되는 것은 아니지만, 고도화된 컴퓨터 처리 능력과 데이터 분석 능력에 의해 개인정보 이외의 데이터를 처리하여 개인에 관한 패턴을 생성할 수 있고, 이러한 패턴만 가지고도 해당 개인을 식별할 수 있으므로 개인에 관한 방대한 양의 데이터(모두 개인정보가 아니라 할지라도)를 보존·분석하는 데에는 특별한 주의와 관심이 필요하다. 빅데이터는 이전에 연결할 수 없었던 정보들 간의 상관관계를 보여줄 수 있고, 이러한 과정에서 개인정보가 생성될 수 있다. 또한 데이터베이스 간의 결합과 분석기술의 이용을 통해 비식별처리된 데이터가 개인정보로 전환될 수 있다. 이러한 맥락에서 개인정보와 비개인정보의 경계에 대한 검토도 요청된다.

정리하면, 빅데이터를 위한 규제체제의 모색에 있어서 개인, 개인의 권익, 개인정보의 개념 등 살아있는 자연인 개인⁶⁸⁾을 부동의 기초로 삼고 있는 현행 규제 체제가 여전히 좋은 유효한 출발점이 될 수 있는지에 대해서 비판적으로 접근할 필요가 있다. 빅데이터 규제에 대하여 현행 규제체제를 적용하거나 개선하는 방식을 선택하든, 새로운 규제체제를 구축하는 방안을 선택하든 상관없이, 개인정보와 관련 정보들 - 예, 민감정보, 재산정보 등 - 을 권리와 의무의 기반으로 삼아 빅데이터를 규제하는 방식에 대한 근본적인 검토가 필요한 시점이다.

덧붙여서, 빅데이터 환경에서는 데이터의 성질이 정태적인 것에서 동태적인 것으로 전환된다는 점에도 주목해야 한다. 빅데이터 처리과정에서 (혹은 데이터 생애주기에 따라) 데이터의 유형이 지속적으로

는 불분명하다.

68) 현행 개인정보 규제체제는 개인과 개인정보처리자의 일대일 관계에 기초하고 있다.

로 변화된다. 현행 개인정보 규제체제는 데이터가 처리과정에서 비교적 정태적일 것으로 전제하고 정보의 유형별로 각기 다른 보호 장치를 적용하는 방식을 취하고 있다. 개인정보, 민감정보, 익명처리된 개인정보, 비식별처리된 정보, 통계정보, 메타데이터 등이 그 예이다. 그러나 빅데이터의 현실은 이와 사뭇 다르다. 빅데이터 처리과정에서 데이터들은 결합, 집적, 비식별처리, 재식별처리 등의 반복적인 순환의 과정을 거치면서 양적 질적 변화를 수없이 거치게 된다. 데이터의 양이 증가하면, 그 처리를 통해 다시 개인을 식별할 수 있는 정보가 되기도 하고, 더 민감한 정보가 되기도 하며, 이렇게 질적으로 변화된 데이터에 다시 비식별조치 등을 적용하여 개인정보 침해에 대한 걱정 없이 통계 분석이나 그룹 프로파일링에 사용할 수 있는 데이터가 되기도 한다. 빅데이터에 대한 새로운 규제적 접근은 이와 같이 데이터 처리에 따른 변화무쌍한 데이터의 동태적 성질에 대한 실제적인 고려가 요청된다.

(3) 빅데이터 활용에 따른 책임과 통제의 문제

빅데이터 규제에 있어서 궁극적인 쟁점은 데이터의 수집·분석·이용에 대한 책임의 귀속주체를 특정하는 문제이다. 빅데이터 환경에서는 데이터 수집·분석·이용 과정에 다수의 다양한 주체들이 참여하게 되고, 이들이 그러한 과정에서 일정한 책임을 공유하거나, 적어도, 관련성을 가지게 된다는 점에 주목할 필요가 있다. 이러한 현상은 책임의 귀속주체를 특정하고, 책임을 배분하며, 적절한 법적 통제를 가하는 문제를 복잡한 난제로 만들 것이다. 구체적으로 빅데이터와 관련하여 최근 나타나고 있는 다음의 세 가지 현상을 법적 책임과 통제의 관점에서 살펴볼 필요가 있다.⁶⁹⁾

첫째, 업무처리의 효율성, 절차의 간소화 등을 목적으로 공공기관 간에 데이터와 데이터베이스의 공유가 증가하고 있다. 이는 한 공공기관이 데이터

를 수집·저장할 수 있는 법적 권한이 제한되어 있음에도 불구하고 필요한 데이터를 수집·저장할 수 있는 더 큰 법적 권한을 가진 다른 공공기관으로부터 많은 정보를 획득할 수 있다는 것을 의미한다. 금융정책을 추진하는 공적 기금관리기관이 국제청의 세무정보나 대법원의 가족관계에 관한 정보를 이용하여 구상권을 행사하는데 사용하는 현황이 하나의 사례이다. 이 경우 누가 규제집행의 책임을 부담하는가의 문제와 법적 통제의 문제가 정당하게 제기된다.

둘째, 민·관·산·학 협력과 협업이 증가하고 있으며, 어떤 경우에는, 그 협력의 범위가 국경을 초월하기도 한다. 국내 공공기관의 전산시스템은 국내 대기업에 의해 구축되었다. 브라질 경찰은 마이크로소프트사와 뉴욕 경찰에 의해 개발된 전산시스템을 사용한다고 한다.⁷⁰⁾ 이 경우, 만약(빅)데이터의 처리·이용과정에서 문제가 발생한다면, 책임의 귀속주체는 누가되는가? 빅데이터 환경에서는 복잡한 데이터 처리 활동의 가치 사슬에서 누가 어떤 역할을 수행하는지를 한눈에 파악하는 일이 점점 더 어려워질 것이다. 또한 대개의 경우 주체의 유형 - 공공기관, 민간기업, 연구기관 등 -에 따라 다른 규제나 또는 법리가 적용되는데 이러한 경우 책임의 정의를 확보하는 문제가 제기될 수 있다.

마지막으로, 국경을 초월하는 데이터 처리 활동이 늘어나고 있다. 문제는 국가마다 개인정보 보호의 수준이 다르기 때문에, 민간 기업들은 강한 규제를 피하기 위하여 상대적으로 규제가 약한 국가에 정착할 가능성이 있다. 그런데 이러한 규범회피의 문제는 정부나 공공기관에서도 발생할 수 있다. 가령, 국내 개인정보 보호법제가 우리 국민들에 대한 국가정보기관의 정보수집권한을 제한한다면, 상대적으로 대한민국 국민을 대상으로 이루어진 정보수집에 대한 법적 제약이 약한 다른 국가의 정보기관이 우리 국민에 대한 데이터를 수집하여 우리 국가정보기관에 그 데이터를 전달할 수 있다.⁷¹⁾ 반대

69) 위의 논문, pp.124-125 참조,

70) 네덜란드 연구보고서, p.41.

71) 실제 네덜란드에서는 이와 관련된 소송이 진행 중 이다.

의 경우도 충분히 가능한 시나리오이다. 정보기관 이외의 다른 국내 공공기관들도 그들에게 적용되는 국내법을 다른 국가의 기관들과 협력하여 효과적으로 회피할 수 있다. 빅데이터 환경에서 이러한 규범회피가, 적어도 이론상으로는, 더욱 수월하게 가능해질 것이다.

빅데이터 환경에서는 빅데이터 수집·처리·이용과정에서 데이터의 성질이 동태적으로 변화하는 것처럼, 책임과 통제 문제도 정태적인 것에서 동태적인 성질로 전화될 것으로 예상되고 관찰된다. 다시 말해서, 정보주체에 대하여 1인의 정보처리자가 데이터를 수집하고 처리하며, 그 주체가 데이터의 주요 처리자로서 개인정보 보호 법령상의 규정과 의무를 이행하여야 하는 정태적인 법의 세상이, 빅데이터의 (이용을 촉진하는) 환경에서 공공기관 간의 데이터 공유 증가, 민관 협력 증가, 국경을 초월한 데이터 이동 증가에 따라 다수의 다양한 주체들에 의한 정보의 수집, 공유, 결합이 이루어지는 동태적인 현실로 변화되어 다가오고 있다. 이와 같은 책임과 통제 문제에 대하여는 1인의 주요 정보처리자를 선택하여 관련 법령상 의무와 책임을 부과하는 방식이나 빅데이터 처리에 관여한 다수의 정보처리자들 간에 책임이 배분·공유·변경되도록 하는 방식 등이 고려될 수 있다.

2. 리스크규제체제로의 빅데이터 규제 패러다임의 전환가능성

(1) 빅데이터 규제에 있어서 리스크규제적 접근방식의 의의

본 연구는 빅데이터의 이용 활성화를 통해 다양한 편익과 혜택을 향유하고 기술 혁신을 촉진하면서 그에 따른 부작용을 가능한 최소화하고 그에 대한 책임 소재를 명확히 할 수 있는 규범적 환경을 조성하기 위해서 현행 개인정보 보호법제의 패러다임과 다른 또는 이를 보완할 수 있는 규제 패러다

임을 모색을 목적으로 한다. 이는 지속적으로 진화하는 빅데이터에 대한 법적 문제를 다른 시각에서 기술적, 사회심리적, 정책적, 경제적 환경에 맞게 재해석하고, 그에 적절한 새로운 규제 패러다임과 법리를 개발하는 시도가 절실히 필요하다는 문제의식의 발로이다. 이상과 같이 빅데이터의 특성, 빅데이터와 관련하여 나타나는 다양한 현상, 관련 법적 쟁점, 현행 개인정보 보호법제에 따른 규제의 한계 등을 검토해 본 결과 빅데이터에 대하여 리스크규제의 방식으로 접근하는 발상의 전환을 시도해 볼 필요가 있겠다는 결론에 이르렀다.⁷²⁾

리스크규제적 접근방식은 빅데이터의 부작용을, 특히 개인정보 및 사생활 침해⁷³⁾ 우려를 사회적으로 큰 편익을 제공하는 고도의 과학기술의 활용에 수반되는 부작용, 곧 ‘리스크’의 문제로 해석하여 이를 관리하는 규제방식이다. 여기에서 리스크란 규범적 개념으로 ‘위해의 개연성과 중대성의 불확실성 속에서 이익(benefits)을 위한 선택에 의하여 발생하는 불확실한 법익침해의 가능성’으로 정의될 수 있다.⁷⁴⁾ 리스크 규제 내지 관리 이론에 기초한 규제 방식은 공동체의 복지, 안전, 과학 기술 혁신, 경제적 효율성, 환경 및 공중보건 등과 같은 근본적인 가치들 간의 조화와 균형(trade-off)을 가능하게 한다. 또한 빅데이터의 활용으로 인한 개인정보 및 사생활 침해의 부작용을 리스크로 해석하여 규

72) 같은 견해로, Alessandro Spina, “Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in EU Data Protection Law?”, EJRR, 2014. 2. pp.248-252; Dennis D. Hirsch, “The Glass House Effect: Big Data, the New Oil and the Power of Analogy”, Maine Law Review, 2014, <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2393792> (최종방문일 2018. 4. 10.); A. Michael Froomkin, “Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements” (November 3, 2015). 2015 U. Ill. L. Rev. 1713 (2015), <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2400736> (최종방문일 2018. 4. 10.).

73) 앞에서 검토한 바와 같이 빅데이터 활용에 따른 부작용은 대개 미흡한 사생활 (및 개인정보) 보호에서 기인하거나 사생활 보호의 문제와 관련이 있으므로, 이하의 논의에서는 사생활 침해를 중심으로 빅데이터 활용의 부작용을 논한다.

74) 리스크의 특징과 개념 정의에 관하여는 윤혜선, 「리스크규제에 관한 公法의 研究 - 食品安全法制를 中心으로」, 서울대학교, 박사학위논문, 2009. 8. 27-19, 61-66면 참조.

Rechtbank's-Gravenhage 23 juli 2014 Zaaknummer: C/09/455237/HA ZA 13-1325, 네덜란드 연구보고서, p.40 참조.

제적 접근을 한다는 것은 개인정보나 개인의 사생활이 단순히 ‘개인의 권리’가 아닌 환경이나 공중 보건 보호와 같은 ‘공동체의 이익’으로 이해될 수 있다는 것을 의미한다.

개인정보 및 사생활 침해를 공동체의 이익으로 이해한다면, 개인정보나 사생활을 침해할 리스크를 야기하는 빅데이터 기술과 그 처리, 관련 기술과 제품 및 서비스 등에 대하여 리스크규제의 이론과 실무에 기반한 접근을 할 수 있다. 즉, 리스크가 발생하는 이유가 고도의 과학기술 혁신에 수반되는 문제이고, 시스템적인 문제이며, 개인에 의해 그 리스크가 적절히 평가·관리될 수 없다면, 이론적으로는 리스크를 평가하여 그 결과에 따라 그에 적절한 규제수단을 선정하여 관리를 하는 방식을 취할 수 있다. 리스크는 빅데이터 활용에 수반되는 위해(사생활 침해)의 성질, 그 위해가 실현된다면 그 크기(리스크의 중대성), 그러한 위해가 실현될 가능성(리스크의 개연성) 등을 종합적으로 고려하여 평가될 수 있다. 나아가 리스크평가 결과를 리스크가 높음 - 중간 - 낮음 - 미미함 등으로 유형화하여, 각각 사전적, 임시적, 사후적, 무규제 등의 적절한 대응적인 규제방안을 마련할 수 있다. 앞에서 검토한 바와 같이 빅데이터의 부작용은 빅데이터의 활용 주체와 목적에 따라 그 중대성이나 개연성이 달라질 수 있는데 리스크규제적 방식은 이러한 기술의 이용 및 규제 환경의 다양성과 동태성을 반영하여 반응적이고 적절한 규제를 가능하게 한다. 또한 빅데이터의 부작용을 공동체의 이익으로 이해한다면, 실무적으로는 다른 분야나 산업에 적용되는 리스크규제의 예를 참고하여, 빅데이터에도 공통적으로 적용될 수 있는 법적 또는 정책적 수단을 선별하여, 빅데이터의 리스크를 사전적 또는 사후적으로 규제하는 방안을 모색할 수 있게 된다.

(2) 리스크규제체제의 전환가능성

빅데이터에 대하여 리스크규제체제를 적용하는 것이 가능한가? 개념적 관점에서, 빅데이터가 야기하는 개인정보 및 사생활 보호의 문제는 특정한 행

위의 결과에 대한 불확실성이 존재한다는 점에서 리스크규제이론에서 정의하는 ‘리스크’와 유사성을 가진다. 위해나 손해의 중대성도 ‘알 수 없고’, 위해나 손해발생의 개연성도 ‘알 수 없는’ 빅데이터의 부작용의 개념과 현실은 인간의 행위의 결과에 대하여 이용할 수 있는 완전한 객관적 과학적 증거가 없는 ‘리스크’ 상황을 의미한다. 또한 빅데이터의 부작용이 ‘리스크’인 이유는, 빅데이터의 특성을 고려할 때, 첫째, 개인정보의 처리에 대하여 개인이 그 고유한 가치나 활용의 의미를 평가할 수 없기 때문이고, 둘째, 빅데이터의 신규성과 가치가 본질적으로 대용량의 데이터와 강력한 분석 수단의 결합에 의해 생성된 결과를 ‘예측할 수 없다’는 점에 있기 때문이다.⁷⁵⁾

제도적 관점에서, 빅데이터의 부작용을 ‘리스크’로, 따라서 새로운 ‘공동체의 이익’으로 볼 수 있다는 주장을 뒷받침하는 근거로는 EU의 GDPR 제35조에 도입된 사전적 개인정보 영향평가 제도를 들 수 있다.⁷⁶⁾ EU GDPR 제35조에 의하여 정보처리자는 개인정보처리를 함에 있어서, 정보처리자의 유형이나 정보처리의 양에 관계없이, 특히 신기술을 사용하거나 또는 정보처리의 성질, 범위, 배경, 목적을 고려하여 자연인의 권리와 자유에 대한 높은 리스크를 초래할 개연성이 있는 경우, 정보를 처리하기 전에, 개인정보 보호에 대하여 예상된 처리작업이 미치는 영향에 대한 평가를 수행하여야 한다.⁷⁷⁾ 우리 「개인정보 보호법」 제33조에도 개인정보 영향평가 제도가 도입되어 있다. 비록 EU GDPR의 사전적 개인정보 영향평가제도의 내용과 수준에 미치지지는 못하지만, 동 규정에 따라 ‘공공기

75) 빅토르 마이어 쾰버거 외, 앞의 책.

76) 같은 견해로는 Alessandro Spina, 앞의 논문, p.248.

77) 나아가 EU GDPR 제35조 제3항은 (i) 자연인의 개인적 사항에 대한 평가가 프로파일링 등 자동화된 처리에 의해 이루어지고, 그 결과에 기초하여 법적 효과를 발생하는 결정이 해당 자연인에 관하여 내려지는 경우 또는 그와 유사한 효과를 가져오는 경우, (ii) 민감정보 등의 대규모 처리 및 범죄행위에 관련된 개인정보의 처리를 하는 경우, 또는 (iii) 공개된 장소에서 대규모의 시스템적인 모니터링을 실시하는 경우 개인정보영향평가를 실시하도록 하고 있다.

관의 장'은 대통령령으로 정하는 기준⁷⁸⁾에 해당하는 개인정보파일의 운용으로 인하여 정보주체의 개인정보 침해가 우려되는 경우에는 그 위험요인의 분석과 개선 사항 도출을 위한 평가를 하고 그 결과를 행정안전부장관에게 제출하여야 한다. 개인정보 영향평가를 하는 경우에는 1. 처리하는 개인정보의 수, 2. 개인정보의 제3자 제공 여부, 3. 정보주체의 권리를 해할 가능성 및 그 위험 정도 및 4. 그 밖에 대통령령으로 정한 사항인 민감정보 또는 고유식별정보의 처리 여부 및 개인정보 보유기간을 고려하여야 한다(동조 제2항, 「개인정보 보호법 시행령」 제36조). 우리 법은 공공기관 외의 개인정보 처리자에 대하여는 개인정보파일 운용으로 인하여 정보주체의 개인정보 침해가 우려되는 경우에 영향평가를 하기 위하여 적극 노력할 것을(제33조 제8항)고 규정하는데 그치고 있다. 이와 같은 개인정보 영향평가제도, 특히, EU GDPR의 사전적 개인정보 영향평가 제도는 신기술인 빅데이터 환경에서 개인정보 및 사생활 보호의 문제를 개인 차원의 사적 집행보다는 공적 규제를 보다 더 필요로 하는 '리스크' 또는 새로운 '공동체의 이익'으로 보려는 EU의 규제 접근방식의 변화를 보여주는 것이라고 할 수 있다.

나아가 이론적 관점에서는 최근 환경 보호와 디지털 프라이버시에 관한 법제도적 발전의 유사성을 논증하는 논문이 발표되고 있다.⁷⁹⁾ 환경법 분야는

리스크 규제, 사전배려의 원칙이 발원한 분야이다. 실무적, 정책적 관점에서는 2014년 미국 대통령 과학기술자문위원회에서 발표한 보고서 “빅데이터와 사생활보호: 기술적 관점(Big Data and Privacy: A Technological Perspective)”에서 사회가 데이터의 생성보다는 데이터의 사용을 통제하는 방식으로 패러다임을 변화시켜야 한다는 점을 강조하였다는 점을 들 수 있다.⁸⁰⁾

물론 화학물질, GMO, 원자력 등이 환경, 공중보건, 안전 등에 야기하는 리스크를 평가하는 것과 같이 빅데이터의 리스크를 평가하는 시스템을 갖추기까지는 제법 많은 시간이 걸릴 수도 있다. 그러나 이와 같은 리스크규제 패러다임으로의 전환을 시도하여, 환경법, 식품법, 원자력법 등과 같은 유사 규제체제에 대한 비교법적 연구를 통해 여 빅데이터에 대한 규제체제를 갖추어 간다면 너무 멀지 않은 미래에 효과적인 빅데이터의 리스크에 대한 평가와 관리가 가능하게 될 것으로 기대한다.

V. 결 론

이상과 같이 본고는 빅데이터 규제의 새로운 패러다임을 모색하기 위하여 빅데이터의 특성을 규제 대상의 관점에서, 특히 새로운 규제 설계의 관점에서 살펴보고, 새로운 규제 패러다임의 필요성 여부를 빅데이터에 대한 현행 개인정보 보호법제의 한계를 중심으로 검토한 후, 새로운 규제 패러다임으로서 리스크규제적 접근가능성을 검토하였다. 사물인터넷, 클라우드컴퓨팅, 빅데이터, 인공지능이 온전히 구현되는 이른바 4차 산업혁명의 환경에서, 외부의 도움 없이, 개인에게 데이터 처리에 수반되는 잠재적 리스크를 자율적으로 통제하도록 요구하는 것은 사실상 불가능한 것을 요구하는 것과 다름

78) 「개인정보 보호법 시행령」 제35조(개인정보 영향평가의 대상) 법 제33조제1항에서 “대통령령으로 정하는 기준에 해당하는 개인정보파일”이란 개인정보를 전자적으로 처리할 수 있는 개인정보파일로서 다음 각 호의 어느 하나에 해당하는 개인정보파일을 말한다.

1. 구축·운용 또는 변경하려는 개인정보파일로서 5만명 이상의 정보주체에 관한 민감정보 또는 고유식별정보의 처리가 수반되는 개인정보파일
2. 구축·운용하고 있는 개인정보파일을 해당 공공기관 내부 또는 외부에서 구축·운용하고 있는 다른 개인정보파일과 연계하려는 경우로서 연계 결과 50만 명 이상의 정보주체에 관한 개인정보가 포함되는 개인정보파일
3. 구축·운용 또는 변경하려는 개인정보파일로서 100만명 이상의 정보주체에 관한 개인정보파일
4. 법 제33조제1항에 따른 개인정보 영향평가(이하 “영향평가”라 한다)를 받은 후에 개인정보 검색체계 등 개인정보파일의 운용체계를 변경하려는 경우 그 개인정보파일. 이 경우 영향평가 대상은 변경된 부분으로 한정한다.

79) Dennis D. Hirsch, 앞의 논문; A. Michael Froomkin, 앞의 논문.
80) 동보고서의 원문은 다음 사이트에서 찾을 수 있다.
<http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf>
(최종방문일 2018. 4. 10.).

없다. 소위 시장 실패라고도 지칭되는⁸¹⁾ 현행 개인 정보 보호법제와 빅데이터 간의 긴장관계는 빅데이터의 개인정보와 사생활 침해 우려를 공적 규제 대상인 ‘리스크로, 또한 개인의 개인정보 자기결정권 및 사생활보호권을 개인의 권리로서 뿐만 아니라 ‘공동체의 이익’으로도 접근할 것을 요청한다.

이러한 관점에서, 다른 분야에서 발전된 리스크 규제체계와 관련 법적정책 수단을 빅데이터와 개인정보 보호 규제에서 어떻게 적용하고 구현할 것인지에 대한 비교법적 연구가 시급하다. 그러한 연구는 빅데이터 활용의 편익을 향유하기 위하여 불가피하게 수반되는 개인의 사적 영역에 대한 간섭을 사회적으로 수용하는 과정에서 서로 상충하는 가치들 간의 조화와 균형을 도모할 수 있는 구체적인 해법을 제시해 줄 수 있을 것이다. 앞에서 언급한 바와 같이, 환경법, 원자력법, 식품법, 의약품법에 이르기까지 현재 리스크규제적 접근이 적용되는 분야의 다양성을 고려할 때 이것은 완전히 새롭거나 매우 어려운 과제는 아니다. 나아가 이와 같은 연구는 인공지능 등 4차 산업혁명을 견인하는 다른 신흥기술들에 대한 규제체제를 구축하는데 있어서도 많은 해안과 실천적 지침을 제시해 줄 수 있을 것이다.

〈참고문헌〉

국내문헌

[단행본]

- 개인정보 보호위원회, 『빅데이터 환경에서 개인정보 보호 강화를 위한 법·제도적 대책방안 연구』, 연구용역보고서, 2012. 12.
- 매일경제기획팀, 서울대 빅데이터센터, 『빅데이터 세상』, 매일경제신문사, 2014.
- 배성환/김동환/곽인호/송용근, 『빅데이터와 SNS시대의 소셜경험전략』, 에이콘출판, 2012.

빅토르 마이어 쾨버거/ 케너세 쿠기어 / 이지연 옮김, 『빅데이터가 만드는 세상』, 21세기북스, 2013.

송민정, 『빅데이터가 만드는 비즈니스 미래지도』, 한스미디어, 2012.

윤혜선, 『리스크규제에 관한 公法的 研究 - 食品 安全法制를 中心으로』, 서울대학교, 박사 학위논문, 2009. 8.

최경진/정준현/구태연/지성우/김도승/성준호, 『빅데이터 환경에서 개인정보보호 강화를 위한 법·제도적 대책 방안 연구』, 개인정보 보호위원회, 2012.

함유근/채승병, 『빅데이터, 경영을 바꾸다』, 삼성경제연구소, 2012.

[논문]

고유흠, “빅데이터와 개인정보보호”, 『지식재산동향』 Vol 21, 한국지식재산연구원, 2014. 12, 70-83면.

김병수, “빅데이터 시대, SNS 사용자의 정보 공유 행태 분석: 프라이버시 계산 이론 관점에서”, 『e-비즈니스연구』 제15권 제1호, 국제 e-비즈니스학회, 2014, 297-315면.

김선경, “빅데이터 시대에 개인화서비스와 프라이버시 패러독스의 고찰 - 사회기술적 관점을 중심으로 -”, 『한국지적정보학회지』 제16권 제2호, 한국지적정보학회, 2014, 193-207면.

김예란, “빅데이터의 문화론적 비판: 미셸 푸코의 생정치 개념을 중심으로”, 『커뮤니케이션이론』 제9권 제3호, 한국언론학회, 2013, 166-204면.

김진영, “빅데이터 분야 규제와 개선방안 검토 - 신정부의 정책방향과 입법과제를 중심으로”, 『IT와 법연구』 제15집, 2017. 8, 157-190면.

손형섭, “개인정보의 보호와 그 이용에 관한 법적 연구”, 『법학연구』 제54집, 한국법학회, 2014, 1-34면.

81) A. Michael. Froomkin, 앞의 논문.

심우민, “스마트 시대의 개인정보보호 입법전략”, 『언론과 법』 제12권 제2호, 한국언론법학회, 2013, 147-174면.

윤혜선, “‘신기술 규제(emerging technologies regulation)’ 연구방법론 개발을 위한 기초 연구 - 기술규제 시 고려요소를 중심으로 -”, 『행정법연구』 제49호, 2017. 6., 103-128면.

양창수, “정보화사회와 프라이버시의 보호 - 사법적 측면을 중심으로 -”, 『인권과 정의』 제175호, 1991, 71-87면.

이재현, “빅데이터와 사회과학: 인식론적, 방법론적 문제들”, 『커뮤니케이션 이론』 제9권 제3호, 한국언론학회, 2013, 127-165면.

조영임, “빅데이터의 이해와 주요 이슈들”, 『한국 지역정보화학회지』 제16권 제3호, 한국지역정보화학회, 2013, 43-65면.

차상욱, “빅데이터(Big Data) 환경과 프라이버시의 보호”, 『IT와 법연구』 제8집, 경북대학교 IT와 법 연구소, 2014, 193-259면.

최대선/김석현/조진만/진승현, “빅데이터 개인정보 위협 분석 기술”, 『정보보호학회지』 제23권 제3호, 한국정보보호학회, 2013, 56-60면.

허성욱, “개인정보유출소송의 현황과 법적 과제 - 법관의 법발견 방법론을 중심으로 -”, 『저스티스』 통권 제110호, 한국법학원, 2009, 302-331면.

외국문헌

[논문]

Froomkin, A. Michael, *Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements*

(November 3, 2015). 2015 U. Ill. L. Rev. 1713 (2015). <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2400736> (최종방문일 2018. 4. 10.).

Hirsch, Dennis D., *The Glass House Effect: Big Data, the New Oil and the Power of Analogy*, Maine Law Review, 2014, <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2393792> (최종방문일 2018. 4. 10.).

Neil, Richardds/Woodrow, Hartzog, *Trusting Big Data Research*, DePaul Law Review, Vol.66:579, 2017.

Spina, Alessandro, *Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in EU Data Protection Law?*, EJRR, 2014. 2. pp.248-252.

Van der Sloot, Bart/Van Schendel, Sascha, *Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study*, 7 (2016) JIPITEC 110.

[자료]

Van der Sloot, Bart/Van Schendel, Sascha, *International and comparative legal study on Big Data*, The Netherlands Scientific Council for Government Policy (WRR) Working Paper 20, 28-04-2016, <<https://www.wrr.nl/publicaties/working-papers/2016/04/28/international-and-comparative-legal-study-on-big-data>> (최종방문일 2018. 4. 1.).

<ABSTRACT>

A Search for New Regulatory Paradigm for Big Data:
Is a shift into the risk-based regulation feasible?

Hye-Sun Yoon*

There are many legal and regulatory issues surrounding Big Data and its use, however, the related legal discussions in Korea are locked up within the boundaries of the Personal Information Protection Act and the related legal system. Regulation of Big Data, hence, calls for a new perspective as well as a new regulatory paradigm and/or legal theory, which can work and be compatible or suitable in fast-changing technical, socio-psychological, policy, and economic environment. Having said that, this article is an attempt to search for a new approach to Big Data regulation, by interpreting the nature of Big Data risk as a side effect of high technology and technology that provides great social benefits.

Hence, this article first examined Big Data as a regulatory subject: it analyzed the concept and definition of Big Data, whether it genuinely exists, whether it is an independent phenomenon, its benefits and risks. Next, this article examined the tension between Big Data and the Personal Information Protection Act, the limitation of the remedies the Act provides for, and of the current regulatory regime under the Act. Then this article explored the possibility and feasibility of shifting the regulatory paradigm into the risk-based regulation. In doing so, this article has first analyzed the crucial issues in designing a new regulatory regime for Big Data: the starting point, i.e. whether the new regime should be based on the concept of ‘personal information’, and the issue of responsibility and control. Finally it has demonstrated that the threats to privacy and personal information can indeed be considered as a risk or new public interest, and thus, a shift to into the risk-based regulation is a real feasibility for Big Data.

Keywords : 빅데이터(Big Data), 규제 패러다임(Regulatory Paradigm), 리스크규제(Risk Regulation), 개인정보 보호법(Personal Information Protection Act), 사생활보호(Privacy)

* Associate Professor, School of Law, Hanyang University.