

기업의 빅데이터(Big Data) 활용과 개인정보의 보호의 조화

곽 관 훈*

목 차

I. 서 론	1. 개인정보보호법의 개요
II. 빅데이터의 의의 및 현황	2. 빅데이터 활용과 개인정보보호법과의 충돌문제
1. 빅데이터의 의의	3. 빅데이터 활용과 개인정보보호의 조화 방안
2. 기업의 빅데이터의 활용 현황과 문제점	
III. 빅데이터 효율적 활용과 개인정보보호의 조화를 위한 방안	IV. 결 론

I. 서 론

ICT(Information & Communication Technology)의 발전으로 스마트시대가 도래함에 따라 우리 일상생활에도 다양한 변화가 나타나고 있다. 특히 중요한 변화 중에 하나는 우리가 생산하는 데이터의 양이 기하급수적으로 증가하고 있다는 것이다. 현대인이 많이 사용하고 있는 SNS의 일종인 트위터(twitter)의 경우 전 세계 1억 명의 월간 이용자들이 이용하고 있으며, 하루 평균 2억 개의 트윗이 발생하고 있다. 또한 페이스북에는 매일 2억 5천만 명이 사진을 업로드하고 있다고 한다.

이처럼 인터넷 및 SNS 등이 보편화되면서 생성되는 디지털 데이터의 양은 2020년에 이르면 현재보다 약 50배 정도 증가할 것으로 예상된다. 이러한 디지털 데이터의 증가는 빅데이터라는 새로운 패러다임을 만들어 내고 있다. 빅데이터는 종래 양적 개념으로서 기존의 체계로는

* 선문대학교 경찰행정법학과 부교수

감당할 수 없을 정도의 거대한 데이터의 집합을 의미하는 용어로 사용되었다. 그러나 최근에는 데이터의 규모가 문제가 아니라 그 데이터를 통해 원하는 정보를 추출하고 이를 통해 의사판단을 할 수 있는 패러다임을 의미하는 개념으로 사용되는 추세이다.

전 세계적으로 빅데이터의 적극적인 활용을 위한 노력이 진행되고 있으며, 우리나라의 경우도 향후 5년 이내에 데이터의 폭발적 증가로 인한 혼돈과 잠재적인 가능성이 공존하는 빅데이터 시대가 도래할 것으로 예상하고 이에 대응하기 위한 방안들을 모색하고 있다.¹⁾ 기업의 빅데이터 활용에 따라 제기되는 법적 쟁점 중 대표적인 것이 개인정보보호와 관련한 문제이다. 기업이 소비자의 구매내역 등 자료를 축적하여, 이를 바탕으로 개인에게 특정 제품을 광고하거나 특정제품의 구매를 추천하는 등의 행위를 하는 경우 이러한 정보들이 개인정보에 해당하는지, 따라서 개인정보보호법의 적용대상이 되는지가 문제가 된다. 이와 관련해서는 빅데이터의 활성화라는 측면도 함께 고려해야 하며, 따라서 지나친 규제가 이루어지지 않도록 합리적인 해결방안을 모색하는 것이 필요하다. 정부는 이러한 문제점을 인식하여 2013년 12월에 ‘빅데이터 개인정보보호 가이드라인’을 제정하였다. 그러나 기존 개인정보보호법과의 관계여서 근본적인 문제들은 해결되지 않고 있으며, 향후 많은 논란이 예상되고 있다. 이하에서는 빅데이터의 활용이라는 측면과 개인정보보호라는 측면을 균형 있게 고려하는 차원에서 그 제도적 대안을 모색해 보고자 한다.

1) 한국정보화진흥원, 새로운 미래를 여는 빅데이터 시대(한국정보화진흥원, 2013), 13쪽.

II. 빅데이터의 의의 및 현황

1. 빅데이터의 의의

1.1 빅데이터의 개념

빅데이터(Big Data)는 Very large DB, Extremely large DB, Extreme data, Total data 등 다양한 용어로 사용되고 있으며²⁾ 명확하게 개념이 정립되지는 않았다. 종래에는 데이터의 양적 측면에서 빅데이터를 정의하는 것이 일반적이었다. 기존의 데이터 베이스 소프트웨어가 수집·저장·관리·분석할 수 있는 범위를 초과한 규모의 데이터를 의미하는 용어로 주로 사용되었다.³⁾ 또한 기술적 측면에서 이러한 데이터로부터 가치 있는 내용을 추출하는 기술을 빅데이터라고 보는 견해도 있다.⁴⁾

그러나 이러한 개념정의들은 빅데이터를 단순히 양적 개념에서 접근하는 것으로 빅데이터가 가진 본질적 의미를 충분히 담아내지 못하고 있다. 빅데이터 기술은 과거 다양한 미디어를 통해 이루어지던 단순 정보전달 방식과 달리, 분석을 통해 의사를 실시간으로 결정하는 수단으로 활용되고 있다는 점을⁵⁾ 고려할 때 양적 측면의 정의로는 충분하지 않기 때문이다.

최근에는 데이터의 형성뿐만 아니라 활용이라는 점을 고려한 개념들

2) 최 성/우성구, “빅데이터의 정의, 활용 및 동향”, 정보처리학회지 제19권 제2호(한국정보처리학회, 2012, 3), 11쪽.

3) McKinsey Global Institute, “Big data: The next frontier for innovation, competition, and productivity”, (2011), 1쪽.

4) John Gantz & David Reinsel, “Extracting Value from Chaos”, IDC IVIEW (2011), 6쪽.

5) 이규철/원희선, “新기술(빅데이터)등장에 따른 경제적 파급효과 및 법(규제)연구”, 정보와 통신 제29권 11호(한국통신학회, 2012, 10), 49쪽.

이 등장하고 있다. 예를 들어 일본 총무성은 빅데이터를 새로운 비즈니스 모델의 개념으로 접근하고 있다. 구체적으로 빅데이터를 ‘사업에 도움이 되는 지식을 도출하는 데이터’로 정의하고 있다. 아울러 빅데이터 비즈니스를 ‘빅데이터를 이용하여 사회·경제적 문제를 해결하고, 업무의 부가가치향상을 행하거나 또는 지원하는 사업’으로 정의하고 있다.⁶⁾ 빅데이터를 데이터의 양적 측면뿐만 아니라 그 데이터를 어떻게 이용하느냐는 질적 측면도 함께 고려하여 파악하고 있다.

우리나라의 경우도 양적인 측면에서 거대한 데이터의 집합을 빅데이터의 개념으로 이해하는 견해도 있으며,⁷⁾ 기술적 측면에서 빅데이터를 정의하는 견해도 있다.⁸⁾ 정부 차원에서 빅데이터에 대한 개념정의가 이루어지고 있는데, 예를 들어 국가정보화전략위원회에서는 ‘대용량 데이터를 활용, 분석하여 가치 있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술’을 빅데이터로 개념정의하고 있기도 하다.⁹⁾

이러한 논의들을 정리해 볼 때, 빅데이터는 양적 측면과 질적 측면을 함께 고려하여 다양한 종류의 데이터를 관리·분석하여 가치 있는 정보를 추출하고 이를 통해 실시간적으로 의사결정을 할 수 있는 시스템 또는 기술을 의미하는 개념으로 이해하는 것이 타당하다.

6) 總務省, 平成24年版 情報通信白書(2012), 153쪽.

7) 이명진/김우주, “빅데이터를 위한 고급분석 기법과 지원 기술”, *Entrue Journal of Information Technology* 제11권 제1호(Entrue정보기술연구소, 2012), 47쪽; 최 성/우성구, 앞의 논문, 12쪽; 박성현, “빅데이터 시대 강 건너 불 아니다”(한국경제 2012.2.3자 참조).

8) 문상일, “빅데이터 환경에서의 금융소비자 개인신용정보 보호를 위한 제언”, *ICT기술의 발전과 경제관련 법제의 쟁점*(한국경제법학회 추계공동학술대회 자료집, 2013), 24쪽.

9) 국가정보화전략위원회, “빅데이터를 활용한 스마트 정부 구현(안)”(국가정보화전략위원회, 2011.10.26).

1.2 빅데이터의 성장배경

빅데이터는 인터넷이 보급됨에 따라 나타나기 시작한 것으로 오늘날 새롭게 등장한 개념이라고 하기는 어렵다. 예를 들어 구글의 검색서비스는 인터넷상의 방대한 웹페이지라는 빅데이터를 활용한 서비스라 할 수 있다. 구글뿐만 아니라 초기 인터넷 포털업체는 대부분 인터넷 빅데이터 검색을 기반으로 성장하였다는 점에서¹⁰⁾ 빅데이터는 새롭게 등장한 개념으로 보기는 어렵다.

다시 말해서 빅데이터가 의미를 갖기 시작한 것은 단순히 데이터양이 증가하였기 때문이 아니라 그 활용 측면에서 변화가 이루어졌기 때문이다. 종래 데이터양의 증가는 정보과잉(information overload) 현상을 초래하여 오히려 혼란을 초래하는 경우가 많았다. 인간이 처리할 수 있는 한계를 초과한 정보의 양적 증가는 오히려 정보처리능력을 저하시키는 부작용이 발생하였던 것이다.¹¹⁾ 하지만 ICT기술의 발전으로 이러한 정보과잉문제를 해결하여 정보의 양적 증가를 유의미하게 활용할 수 있도록 하였다. 이러한 기술발전에 따라 2011년 이후 빅데이터라는 용어는 기술뿐만 아니라 사회, 문화, 정치 등을 아우르는 새로운 패러다임으로 부각되기 시작한 것이다.¹²⁾ 이처럼 빅데이터가 급성장하고 그 의미도 변화하기 시작한 배경으로는 데이터양의 급격한 증가 및 이러한 데이터를 분석하는 ICT기술의 발전을 들 수 있다.

10) 백옥인, “빅데이터의 형성과 전유체계 비판”, 동향과 전망 제87호(한국사회과학연구소, 2013), 308쪽.

11) 이환수/임동원/조향정, “빅데이터 시대의 개인정보 과잉이 사용자 저항에 미치는 영향”, 지능정보연구 제19권 제1호(한국지능정보시스템학회, 2013, 3), 127쪽.

12) 장영재, “빅데이터의 활용 -경영을 과학으로-”, 상장협 춘계회(한국상장회사협의회, 2013), 109쪽.

1.2.1 데이터양의 증가 및 활용가치의 변화

인터넷의 보편화, 스마트폰의 보급에 따라 매일 생성되는 데이터의 양은 기하급수적으로 증가하고 있다. 예를 들어 월마트의 경우 시간당 100만 건 이상의 거래기록이 저장되고 있으며, 트위터의 경우 2011년 기준으로 매일 1억 1,000만 건의 트윗이 발생하고 있다.¹³⁾ 이처럼 매일 생성되는 정보의 양은 기하급수적으로 증가하고 있다. 인류가 탄생하고 2003년까지 쌓아 온 데이터의 총량은 약 5엑사 바이트 정도인데, 2004년부터는 이 정도 규모의 데이터를 생성하는데 불과 이들의 시간밖에 걸리지 않는다고 한다.¹⁴⁾ 지금 추세대로 정보가 계속 증가하면 2020년에는 정보의 양이 지금보다 약 50배 이상 증가할 것으로 예상되고 있다.¹⁵⁾ 또한, 데이터의 성격도 과거 텍스트 위주의 데이터에서 그림, 동영상, 음성 위주의 비정형 데이터의 비중이 증가하고 있다.¹⁶⁾

이처럼 데이터양이 급증한 주요 원인 중의 하나로 M2M(Machine to Machine)의 성장을 들 수 있다. M2M이란 기계 및 장치에 탑재된 센서로부터 정보를 자율적으로 통지하는 기능을 가지고, 그 정보를 가지고 감시 및 제어 등의 서비스를 제공하는 시스템을 의미한다.¹⁷⁾ 이러한 M2M의 성장에 따라 다종·다양한 정보들이 자동으로 축적될 수 있으며, 이는 빅데이터를 형성하는 데 중요한 역할을 하고 있다.¹⁸⁾

ICT기술의 발전에 따른 변화는 기존에도 존재하였다. 원격지에서의 대화를 가능하게 하였던 전화의 발명은 우리 인류에게 커다란 변화를 가져왔으며, 그 밖에도 전자메일, 화상통화, 전자문서교환, 스마트 폰

13) 최경진 외, 빅데이터 환경에서 개인정보보호 강화를 위한 법·제도적 대책 방안 연구(개인정보보호위원회 연구용역보고서, 2012), 7쪽.

14) 장영재, 앞의 논문, 109쪽.

15) 안창원/황승구, “빅데이터 기술과 주요이슈”, 정보과학지 제30권 제6호(한국정보과학회, 2011), 10쪽.

16) 최 성/우성구, 앞의 논문, 13쪽.

17) 奥室 慈, “M2M とビッグデータ時代の到来”, KIIS 144호(2012), 3쪽.

18) 風間博之, “M2M クラウド時代のプラットフォームサービス”, KIIS 144호(2012), 7-8쪽.

등의 등장도 많은 변화를 가져왔다. 그럼에도 불구하고 오늘날 빅데이터의 성장이 특별한 이유는 과거 다양한 미디어의 경우 단순히 정보를 전달하는데 그쳤던 반면에, 빅데이터는 분석을 통해 실시간으로 의사를 결정하는 수단으로 활용되고 있기 때문이다.¹⁹⁾

종래 데이터양이 부족했던 시기에는 제한된 정보를 통해 전체를 유추하거나 한정된 데이터를 바탕으로 전체를 판단하고자 노력하는 ‘통계’라는 수학적 방식들이 이용되었다. 통계는 본질적으로 데이터가 충분하지 않다는 전제 하에서, 부족한 데이터를 바탕으로 전체를 파악하고자 하는 방식이었던 것이다.²⁰⁾ 또한, 이러한 통계를 통해서 얻는 정보는 주로 과거의 트렌드에 대한 분석이었다. 즉, 데이터가 생성되는 시점과 데이터를 취합하여 분석하는 시점 간에 시차가 존재함에 따라 데이터를 통해 얻을 수 있는 것은 과거에 대한 분석이었던 것이다. 하지만 빅데이터의 경우 충분한 정보를 손쉽게 얻을 수 있기 때문에 데이터 자체가 중요한 것이 아니라 데이터를 통해 어떤 결과를 도출하고 어떤 의사결정을 하는가가 중요한 문제가 되고 있다.²¹⁾ 즉, 데이터 분석을 통해 실시간으로 의사결정을 내릴 수 있다는 점이 최근 빅데이터가 중요한 의미를 갖게 된 가장 큰 이유라고 할 수 있다.

1.2.2 ICT의 기술의 진화

살펴본 바와 같이 데이터는 그 자체로 중요한 것이 아니라 그 데이터를 어떻게 분석하는가가 중요한 의미를 갖는다. 이와 관련해서 중요한 의미를 갖는 것이 바로 ‘애널리틱스(Analytics)’이다. 애널리틱스는 종래 일반적인 비즈니스 분석과 달리 복잡한 연산을 수학적 최적화나 고도의 알고리즘을 통해 실시간으로 분석해 그 결과를 도출하는 의미로

19) 이규철/원희선, 앞의 논문, 49쪽.

20) 장영재, 앞의 논문, 113쪽.

21) 위의 논문, 114-115쪽.

통용되고 있다.²²⁾ 하지만 정형화된 개념은 아니며 빅데이터로부터 의미있는 가치를 얻어내기 위한 다양한 접근방법을 의미한다.²³⁾

오늘날 기업들은 다양한 애널리틱스를 가지고 주어진 데이터를 분석하여 의사결정을 내릴 수 있는 기반을 구축하고 있다. 종래 많은 양의 데이터를 분석하고 활용할 수 있는 능력을 갖춘 기업은 구글 등 데이터 전문기업에 한정되었다. 그러나 ICT의 발전에 따라 데이터 전문 기업이 아닌 일반 기업들도 다양한 애널리틱스 기술을 활용할 수 있으며, 이에 따라 빅데이터의 분석을 통해 의미있는 결과를 도출하는 것이 가능하게 되었다.²⁴⁾

1.3 빅데이터의 특성

기존 데이터와 빅데이터의 차이점에 대해서는 일반적으로 ① 데이터의 규모(Volume), ② 데이터 처리 및 분석의 속도(Velocity), ③ 데이터의 다양성(Variety)의 3Vs를 제시하고 있다.²⁵⁾

종래 빅데이터의 개념 자체를 양적 팽창에서 찾았던 것처럼 빅데이터가 가진 가장 큰 특징이 데이터의 규모가 방대하다는 점은 명확하다. 다만, 단순히 데이터의 양적 팽창만을 의미하는 것이 아니라 종래 많은 시간이 요구되었던 데이터 처리와 분석이 짧은 시간 내에 이루어짐에 따라 실시간으로 의사결정이 가능하게 되었다는 점도 중요한 특징이 되고 있다.²⁶⁾

22) 위의 논문, 111쪽.

23) 김근태, “빅데이터 분석을 위한 기업의 Big Analytics 환경 변화”, 정보처리학회지 제19권 2호(한국정보처리학회, 2012), 71쪽.

24) 장영재, 앞의 논문, 111쪽.

25) 이광석, “정보재난과 빅데이터 위험 정보사회”, 문화과학(문화과학사, 2012), 152쪽; 최경진 외, 앞의 책, 10쪽 이하 참조. 이에 대해 위 세 가지 요소에 데이터의 가치(Value)까지 포함한 4V를 빅데이터의 특징이라고 하는 견해도 있다; 김근태, 앞의 논문, 70쪽.

또한 데이터의 유형도 정형화된 데이터뿐만 아니라 비정형화된 데이터들로 다양화되고 있다. 예를 들어 XML, HTML 등과 같은 반정형 데이터를 이용한 분석뿐만 아니라 카톡메시지, 페이스북 코멘트 등 SNS 데이터나 로그파일 같은 비정형 데이터 등의 다양한 정보가 처리되고 있다는 점이 빅데이터의 중요한 특징 중의 하나라고 할 수 있다.²⁷⁾

이처럼 빅데이터는 데이터 규모나 다양성, 처리 및 분석의 속도 등에서 기존의 데이터와는 많은 차이점을 보이고 있다. 아울러 앞서 설명한 바와 같이 애널리틱스의 진화에 따라 종전에는 전문가만이 분석할 수 있었던 것을 일반기업들도 쉽게 활용할 수 있게 되었다는 점도 빅데이터의 중요한 특징 중 하나이다.

2. 기업의 빅데이터의 활용 현황과 문제점

2.1 빅데이터 활용 현황

2.1.1 개요

스마트시대를 맞이하여 국가적 차원에서 전자정부에 관심이 증가하고 있으며,²⁸⁾ 빅데이터의 활용도 향후 지속적으로 증가할 것으로 예상되고 있다.²⁹⁾ 그러나 무엇보다도 중요한 의미를 갖는 것은 기업들

26) 이에 관한 상세한 내용은 최경진 외, 위의 책, 10쪽 이하 참조.

27) 성준호, “빅데이터 환경에서 개인정보보호에 관한 법적 검토”, 법학연구 제21권 제2호(경상대학교 법학연구소, 2013), 311쪽.

28) 공공영역에서도 빅데이터의 활용을 위한 노력들이 지속적으로 전개되고 있다. 우리 정부는 2011년에 ‘스마트 전자정부 계획’과 ‘모바일 전자정부 서비스 중장기 추진계획’을 수립하여 스마트 시대로의 전환을 추진하고 있다. 또한 클라우드의 도입 및 확산을 위해 ‘클라우드 컴퓨팅 확산 및 경쟁력 강화 전략’을 마련하기도 하였다.

29) 정부는 국가 차원에서 빅데이터의 활용하기 위한 다양한 방안들을 모색하고 있다. 특히, 의료분야, 세제분야, 교통분야 등에 있어서 빅데이터의 활용은 사회적 비용의

의 빅데이터 활용이라고 할 수 있다. 종래부터 기업에 있어서 주어진 데이터를 어떻게 활용하는가의 문제는 기업 경쟁력과 직결되는 문제였다.³⁰⁾ 예를 들어 의류회사가 기존 데이터 분석을 통해 향후 유행할 디자인 등을 예측하는 것은 기업의 존립과 직접 관련되는 문제이기 때문이다.

이러한 점을 고려할 때 엄청난 규모로 생산되는 빅데이터 시대의 도래는 기업에 있어서는 위기이자 새로운 기회라고 할 수 있다. 빅데이터의 활용을 통해 경영을 혁신하고 생산성 향상, 원가 절감 및 미래 트렌드 예측을 통한 신상품 개발 등에 성공하는 기업은 향후 지속가능한 발전을 할 수 있을 것이다. 그러나 늘어난 데이터 저장공간과 보다 정교해진 분석 기술은 더 많은 데이터의 수입 및 분석을 유도하고 있으며 이러한 흐름에 따라가지 못하는 기업은 도태될 수밖에 없을 것이다.³¹⁾ 최근 빅데이터를 활용하여 많은 성과를 거두고 있는 기업들이 점차 증가하고 있다. 이하에서는 주요 사례에 대해 살펴보기로 한다.

2.1.2 빅데이터 활용의 대표적 사례

2.1.2.1 Catalina Marketing 사례

미국의 Catalina Marketing의 경우 고객이 대형 할인마트 매장 계산대

감소는 물론 국가발전에 크게 기여할 수 있을 것으로 예상하고 있다. 본문에서는 기업을 중심으로 논의를 함에 따라 공공부문에 대한 논의는 생략한다. 이에 대한 상세한 내용은 한국정보화진흥원, 앞의 책, 190쪽 이하 참조.

30) 미국의 MIT공대에서 발행하는 잡지인 ‘MIT Sloan Management Review’가 IBM사의 연구조직인 ‘IBM Institute for Business Value’와 공동으로 100여 개국 기업 경영간부, 매니저, 분석담당자 약 3천 명을 대상으로 조사한 결과에 따르면 업계 상위인 기업은 하위인 기업과 비교할 때 데이터 분석을 5배 이상 활용하는 것으로 나타나고 있다. 城全眞琴, “日本におけるビックデータの現想と課題”, IT Solutions Frontier 29권 3호(野村總合研究所, 2012), 8-9쪽.

31) 이환수 외, “빅데이터 시대의 개인정보 과잉 인식이 사용자 저항에 미치는 영향”, 한국지능정보시스템학회 학술대회 자료집(2012년 가을), 36쪽.

에서 결제를 하는 순간 고객의 구매이력 데이터가 분석되어 이 고객을 위한 맞춤형 쿠폰이 인쇄된 영수증이 출력된다. 고객이 동일 제품을 구매하여도 구매이력에 따라 각각 다른 쿠폰이 출력되며, 이를 통해 추가적인 구매를 유도하는 것이다. 이러한 결과를 얻기 위해 Catalina Marketing은 DW 어플라이언스에 High Performance Computing 분석 기술을 접목하여 2.5 Peta Bytes의 EDW로부터 고객의 구매 성향을 실시간으로 예측하였다. 동 기술의 이용에 따라 종래 4.5시간 걸리던 예측모델에 대한 스코어링 시간도 불과 60초로 단축시킬 수 있었다.³²⁾

2.1.2.2 아마존닷컴(Amazon.com)의 사례

아마존닷컴은 회원들 개개인의 구매 패턴을 파악한 후 자동화된 알고리즘으로 개인에 맞는 쿠폰을 생성하여 고객에게 전달하는 서비스를 제공하였다. 예를 들어 유아용 세발자전거를 구매한 후 얼마 지나지 않아 안장벨트를 구매한다는 패턴을 구매 데이터를 통해 파악한 후, 세발자전거를 산 후 아직 안장벨트를 구매하지 않은 고객에게 안장벨트 할인쿠폰을 제공하는 것이다.³³⁾

또한 아마존을 통해 도서를 구입한 고객의 검색어와 도서 구입 패턴 분석을 통해 이전에 특정 도서를 구입한 사람이 다음에 어떤 도서를 구입했는지 추천해주는 서비스를 제공하고 있다. 아마존 고객의 30% 정도가 추천 알고리즘을 통해 물품을 구입하는 것으로 알려져 있다.³⁴⁾

2.1.2.3 자라(Zara)의 사례

자라는 패션업체로 빅데이터를 활용한 모범사례 중 하나로 평가되고 있다. 보통 패션업체의 경우 한 시즌 판매할 상품들을 6개월에서 길게는 1년 전부터 기획하여 생산한 후 매장에 진열한다. 이에 대해 자라는

32) 김근태, 앞의 논문, 71쪽.

33) 장영재, 앞의 논문, 114쪽.

34) 최경진외, 앞의 책, 14쪽.

불과 몇 주 내에 기획, 디자인, 생산, 매장 진열까지 완료하여 시의적절한 제품을 제공하는 것이 가능하다. 이것이 가능한 것은 자라만의 애널리틱스 방식을 개발하여 전 세계 매장에서 판매되는 데이트를 수집, 분석하여 상품 수요 예측, 각 매장별 적정 재고 산출, 그리고 최적화된 상품별 가격 결정 등에 활용하고 있기 때문이다.³⁵⁾

자라의 성공비결은 빅데이터 분석을 통해 실시간으로 필요한 의사결정을 내리는 방식을 취하였다는 점이다. 다른 패션업체의 경우 상품 판매에 관련한 데이터를 통해 1년 후의 패션 트렌드를 예측하였지만, 자라의 경우 현장에서 실시간으로 유입되는 방대한 데이터를 통해서 현재의 트렌드를 분석하고 대응하였으며, 이것이 자라의 성공비결로 평가되고 있다.³⁶⁾

2.2 빅데이터 활용에 따른 문제점

기업이 빅데이터를 활용함에 있어 발생할 가능성이 높은 법적 문제는 개인정보보호와 관련한 문제이다. 위의 사례에서 본 바와 같이 기업이 소비자의 기호를 파악하고, 소비자가 원하는 물건을 소비자보다도 먼저 파악하여 마케팅에 활용하는 것이 빅데이터가 갖는 가장 큰 의미이다. 기업 입장에서는 새로운 비즈니스 모델로서 큰 의미를 갖게 된다. 그러나 소비자 개인의 입장에서는 자신의 의사와 관계없이 기업에 개인정보를 이용함에 따라 정보유출의 문제가 발생할 수 있다.

빅데이터 환경하에서는 지금까지의 개인정보 유출과는 또 다른 측면에서 문제가 발생할 가능성이 높다. 앞서 살펴본 바와 같이 빅데이터는 방대한 데이터를 통해 새로운 정보들을 추출할 수 있다. 이에 따라 정보주체는 지금보다도 개인 프라이버시가 침해될 위험성을 우려하게 되고, 이러

35) 장영재, 앞의 논문, 116쪽.

36) 위의 논문, 116-117쪽.

한 우려는 해당 시스템 사용 자체의 사용을 꺼리는 원인이 될 수 있다. 즉, 개인정보의 제공 자체에 상당한 거부감을 가질 수 있으며,³⁷⁾ 이는 결과적으로 빅데이터를 의미 없게 만드는 요인이 될 수 있는 것이다.

이처럼 빅데이터는 이중적 측면을 가지고 있다. 기업의 국제경쟁력 강화를 위해서는 빅데이터를 적극 활용할 필요가 있다. 빅데이터 시장은 그 규모가 급격하게 증가하고 있으며,³⁸⁾ 기업의 새로운 시장으로서 중요한 의미를 가지고 있다.³⁹⁾ 이러한 상황에서 개인정보보호가 지나치게 강조되는 경우 국내시장이 위축될 우려가 지적되고 있다. 반면에 빅데이터 활용으로 인해 개인정보가 침해된다는 우려 또한 계속 지적되고 있다.⁴⁰⁾

빅데이터의 효율적 활용을 위해서는 개인정보가 적절하게 보호되면서도 한편으로는 이러한 정보들이 자유롭게 활용될 수 있는 환경이 갖추어지는 것이 필요하다. 즉, 법적 측면에서 개인정보 보호와 빅데이터 효율적 활용 사이에 적절한 접점을 찾는 것이 중요한 과제가 되고 있다.

37) 실제 빅데이터 환경하에서 증가된 프라이버시 염려수준은 사용자들의 시스템 사용에 대한 부정적 태도인 저항적 태도를 형성하여 궁극적으로 시스템 사용을 중단할 수 있다는 실증적인 연구가 있다. 이환수 외, 앞의 논문, 125-134쪽 참조.

38) 시장조사기관인 IDC에 따르면 세계 빅데이터 시장은 2012년 68억 달러에서 2015년 169억 달러로 성장할 것으로 예상하고 있다.

39) 기업뿐만 아니라 공공부문에서도 빅데이터의 적극적인 활용을 위한 노력을 경주하고 있다. 최근 정부는 2017년까지 97개 빅데이터 활용사업을 추진하면서 올해 예산도 지난해 360억 원보다 증가한 460억 원을 책정하였다. 올해는 25개 신규사업을 추진하고 있으며 2013년에 완료된 여성가족부의 ‘위기청소년 조기경보’, 경찰청의 ‘범죄확률 사전예측’, 통계청 ‘일일물가분석’ 등 주요과제들의 서비스도 본격적으로 제공될 예정이다. 또한 ‘빅데이터 공통기반 플랫폼’을 통해 행정기관들이 공개된 공공과 민간의 데이터를 자유롭게 수집, 분석, 공유할 수 있도록 지원할 예정이다. 국민일보 2014년 1월 8일자 16면 참조.

40) 문상일, 앞의 논문, 28-29쪽.

Ⅲ. 빅데이터 효율적 활용과 개인정보보호의 조화를 위한 방안

1. 개인정보보호법의 개요

2011년 3월에 제정된 ‘개인정보보호법’은 기존의 ‘공공기관의 개인정보보호에 관한 법률’과 ‘정보통신망 이용촉진 및 정보보호 등에 관한 법률’의 일부 조항 등을 포괄하여 개인정보보호에 관한 기본법으로서 제정되었다. 종래 다양한 법률에서 산재되어 있던 개인정보보호에 관한 규정을 종합적으로 정리하여 보호의 사각지대를 해소하기 위해 제정된 법이다.

개인정보보호법의 보호대상인 ‘개인정보’란 살아 있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보를 말하며, 해당 정보만으로는 특정개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보를 포함하는 개념으로 이해하고 있다(제2조 1호). 즉, 개인을 식별할 수 있는 정보가 개인정보에 해당되는데, 주의할 점은 한 가지 정보만으로는 특정인을 알아볼 수 없다고 하여도 정보의 결합 또는 조합을 통하여 특정개인을 구분하거나 구별할 수 있으면 개인정보에 해당한다.⁴¹⁾

개인정보보호법은 정보주체가 자신의 개인정보가 언제, 어디서, 어떻게, 어느 범위까지 수집·이용·제공되는지에 대해 스스로 판단·결정할 수 있는 개인정보 자기결정권을 갖고 있다는 점을 전제로 하고 있다.⁴²⁾ 따라서 개인정보를 처리하는 자는 개인정보보호에 관한 기본 원칙(제3조)을 준수해야 하며, 개인정보를 수집함에 있어 개인정보의

41) 행정안전부, 개인정보 보호법령 및 지침·고시 해설(행정안전부, 2011.12), 8쪽.

42) 김운석, “개인정보보호 2.0시대의 개인정보보호법 개관”, 법학연구 제22권 제2호 (충남대학교 법학연구소, 2011.12), 13쪽.

수집·이용 목적 등을 알리고 정보주체의 동의를 받는 것이 원칙이다(제15조 1항 1호, 2항). 또한 개인정보를 수집함에 있어서도 그 목적에 필요한 최소한의 정보를 수집해야 하는 등(16조 1항) 엄격한 제약이 따르고 있다.

살펴본 바와 같이 개인정보보호법의 경우 개인정보의 수집 등의 경우에 정보주체의 동의를 얻는 것을 원칙으로 하고 있다. 정보주체의 동의를 얻는 방법은 크게 opt-in 방식과 opt-out 방식으로 나눌 수 있다. opt-in 방식은 개인정보를 수집하기 위해서는 사전에 정보주체에게 처리 목적을 고지하고 정보주체의 명시적 동의(affirmative consent)를 받아야 하는 방식을 말한다. 반면에 opt-out 방식은 개인정보의 수집, 처리에 관하여 우편, 전자우편 등을 통해 정보주체에게 알리고 이에 대해 정보주체가 공식적으로 이의를 제기하지 않으면 개인정보 활용이 허용되는 방식이다. 미국의 경우 opt-out 방식을 채택하고 있는 것에 비해 유럽과 우리나라는 opt-in 방식을 채택하고 있다.⁴³⁾

2. 빅데이터 활용과 개인정보보호법과의 충돌문제

2.1 개인정보에의 해당성

개인정보보호법에 의하면 다른 정보와 결합하여 개인을 식별할 수 있는 경우 개인정보로 보고, 이를 수집하거나 이용하기 위해서는 정보주체에게 사전에 동의를 얻어야 한다. 빅데이터의 경우도 다른 정보와 결합하여 개인을 식별할 수 있는 경우 개인정보에 해당되어 이와 같은 규제를 받을 수 있다. 또한 기업이 빅데이터를 활용하여 개인의 식별정보와 행위정보를 결합하여 타겟 마케팅을 운용하거나, 자사의 서비스를

43) 이창범, “개인정보보호법제 관점에 본 빅데이터의 활용과 보호 방안”, 법학논총 제37권 제1호(단국대학교 법학연구소, 2013), 519-521쪽.

대가로 수집된 데이터를 기반으로 광고에 응용하거나 데이터 자체를 판매하기도 한다. 이 경우 개인정보 보호의 문제는 물론 프라이버시의 문제가 발생할 수도 있다.⁴⁴⁾

물론 빅데이터의 경우 기존의 개인정보와는 여러 가지 면에서 차이가 있다. 기존의 경우 기업이 소비자의 정보를 수집하고, 소비자 등도 이러한 사실을 인식한 상태에서 개인정보를 제공하였다. 그러나 빅데이터의 경우는 소비자 등이 기업에게 제공한다고 인식하지 않는 상태에서 데이터가 형성되는 경우가 많다. 예를 들어 아마존닷컴에서 구매한 내역 등은 소비자가 기업에게 의식적으로 제공하는 정보에 해당되지 않는다. 또한 빅데이터의 상당수는 클라우드 서비스를 통해 수집됨에 따라 정보제공자의 의사와 관계없이 수집되는 경우가 많다.⁴⁵⁾ 이렇게 수집된 정보를 기업이 특정한 ‘애널리틱스(Analytics)’를 이용하여 분석해서 유용한 정보를 추출하는 것은 기존의 ‘개인정보’와는 분명한 차이점을 가지고 있다.

그러나 일반적으로는 개인을 식별할 수 있는 정보를 개인정보로 이해함에 따라 사용자들이 스스로 공개한 좋아하는 곳, 좋아하는 음식 등에 관한 정보도 다른 정보와 결합하여 특정인을 식별할 수 있다면 개인정보 침해에 해당한다고 볼 가능성이 높다. 이에 따라 기업들의 빅데이터 활용이 위축되고 있는 상황이다.⁴⁶⁾

44) 신영진, “공공분야의 빅데이터 추진과 개인정보보호에 관한 연구: 개인정보 보호법의 주요 내용과 개선과제를 중심으로”, Internet and Information Security 제3권 제3호(한국인터넷진흥원, 2012), 93쪽; 백옥인, 앞의 논문, 306쪽; 이규철/원희선, 앞의 논문, 48쪽; 신영진, 앞의 논문, 94쪽 이하 참조.

45) 신영진, 위의 논문, 97쪽.

46) 실제로 SK텔레콤의 경우 2013년 10월 가입자들의 통화내역과 멤버십 포인트 이용 현황을 토대로 만든 빅데이터를 일반에 공개하였다. 무슨 요일에 중국집, 피자 배달 주문이 물리는지, 서울 어느 지역에서 대리운전, 콜택시 이용이 많은지 등 10종의 정보를 공개하였다. 하지만 당초 SK텔레콤이 공개하려 했던 데이터는 100종이 넘었다. 그러나 현행법상 개인정보 유출로 제재를 받을 가능성이 있어 공개범위를 대폭 축소한 것이다. 중앙일보 2013년 11월 10일자 참조.

2.2 정보주체의 동의 방법

빅데이터가 개인정보에 해당한다면 이를 수집·이용하기 위해서는 원칙적으로 정보주체의 동의를 얻어야 한다(제15조)⁴⁷⁾ 동의를 얻은 경우에도 필요한 범위 내에서 최소한의 정보만 수집해야 하며(제16조), 수집목적을 넘어 이용하고자 하는 경우 정보주체로부터 별도의 동의를 얻어야 한다(제18조 2항). 또한 보유기간이 경과하거나 목적을 달성한 개인정보는 지체 없이 파기하여야 한다(제21조).

앞서 살펴본 바와 같이 빅데이터는 단순한 데이터의 집합체를 의미하는 것이 아니라, 해당 데이터로부터 유의미한 정보를 추출하여 활용하는 것을 말한다. 예를 들어 소비자가 쇼핑몰에서 물건을 구매한 내역 자체가 중요한 것이 아니라, 그 정보를 통해 특정 소비자가 향후 어떠한 물건을 구매할 것인지를 예측하고 이를 마케팅에 활용하는 것이 중요한 것이다. 따라서 단순히 개인정보를 수집하는 것에 그치지 않고 이를 분석하고 활용하여야 한다. 또한, 개인정보 수집당시에는 생각하지 못했던 목적으로 활용할 가능성도 크다.

이러한 빅데이터에 대해 개인정보보호법을 그대로 적용한다면 정보를 수집하는 단계에서도 정보주체의 동의를 얻어야 할 뿐만 아니라, 이를 활용하여 분석하는 모든 단계마다 동의를 얻어야 하는 문제가

47) 정보주체의 동의를 얻는 경우 외에도 ① 법률에 특별한 규정이 있거나 법령상 의무를 준수하기 위하여 불가피한 경우, ② 공공기관이 법령 등에서 정하는 소관 업무의 수행을 위하여 불가피한 경우, ③ 정보주체와의 계약체결 및 이행을 위하여 불가피한 경우, ④ 정보주체 또는 그 법정대리인이 의사표시를 할 수 없는 상태에 있거나 주소불명 등으로 사전 동의를 받을 수 없는 경우로서 명백히 정보주체 또는 제3자의 급박한 생명, 신체, 재산의 이익을 위하여 필요하다고 인정되는 경우, ⑤ 개인정보 처리자의 정당한 이익을 달성하기 위하여 필요한 경우로서 명백하게 정보주체의 권리보다 우선하는 경우 등에는 개인정보를 수집, 이용할 수 있다. 그러나 빅데이터의 경우에는 이러한 사유에 해당하는 경우를 예상하기 어려우며, 따라서 정보주체의 동의를 받아 정보를 수집하는 경우만을 염두하고 서술한다.

발생한다. 물론 개인정보에 해당하지 않는 경우는 해당되지 않을 것이다. 그러나 무엇이 개인정보에 해당하는지가 명확하지 않은 상황에서 정보주체가 정보제공 자체를 꺼리거나,⁴⁸⁾ 개인정보유출을 주장하여 법적 분쟁이 발생할 가능성이 매우 높다.

3. 빅데이터 활용과 개인정보보호의 조화 방안

3.1 입법적 정비의 필요성

빅데이터의 활용과 개인정보보호법의 충돌문제가 발생하자 정부는 이를 해결하기 위한 방안으로서 ‘빅데이터 개인정보 보호 가이드라인’의 제정을 추진하고 있다.⁴⁹⁾ 동 가이드라인은 개인이 제한 없이 일반 공중에게 공개한 부호나 문자, 음성 등 개인식별이 가능한 정보를 ‘공개된 정보’로 규정하고, 공개된 정보를 수집, 활용하거나 제3자에게 전달할 때 정보주체의 동의를 받지 않아도 되도록 하고 있다. 또한 정보통신 이용내역을 수집하는 경우 및 수집된 정보를 활용해 새로운 정보를 생성하는 경우에도 정보주체의 동의를 요하지 않도록 하고 있다.⁵⁰⁾

정부가 이러한 가이드라인을 제정하고자 하는 이유는 개인정보보호

48) 실제로 빅데이터의 활용이 증가하면서 개인들이 개인정보 공개를 꺼리는 현상이 나타나고 있다. 시장조사업체 Ernst&Young이 일반소비자와 기업인원을 대상으로 진행한 ‘빅데이터 반발(BigData Backlash)’ 보고서에 따르면 소비자들은 기업들의 개인정보 수집을 우려하고 있으며, 앞으로 공개수준을 낮출 것으로 보인다. 응답자 중 55%가 이전보다 자신의 개인정보를 적게 공개하고 있으며, 향후 자신의 개인정보에 대한 기업들의 접근을 제한하겠다고 하였으며, 특히 개인정보가 본인이 모르는 제3자가 활용하는 것에 대해 민감하게 생각한다고 한다. 디지털 타임스, 2014년 1월 3일자 6면 참조.

49) 방송통신위원회는 ‘빅데이터 개인정보 보호 가이드라인(안)’을 마련하여 2013년 12월 18일에 공청회를 가졌으며, 각계 의견을 수렴하고 있다.

50) 가이드라인은 공개된 개인정보는 수집 시 정보주체 동의 없이 가능하며, 정보통신 이용내역을 수집하는 경우나 수집된 정보를 활용해 경향신문 2013년 12월 30일자 참조.

법상의 규정이 지나치게 엄격하여 빅데이터 활용에 저해가 된다는 판단 때문이다. 그러나 가이드라인이 개인정보를 침해할 우려가 있다는 주장들이 제기되고 있다.⁵¹⁾ 더 큰 문제는 가이드라인이 개인정보보호법 상 규정과 상충되고 있다는 점이다. 예를 들어 가이드라인의 경우 공개된 개인정보 수집 시 정보주체의 동의를 요하지 않는데, 이는 개인정보 수집 시 정보주체의 동의를 얻도록 하고 있는 개인정보보호법 규정과 배치되는 문제가 발생하고 있다. 결국 가이드라인의 제정을 통해 빅데이터 활용과 관련한 문제를 해결하기는 어려우며, 입법적인 정비를 통해 해결되어야 할 것이다.

3.2 구체적 개선방안

3.2.1 개인정보의 구분

빅데이터와 관련하여 개인정보보호법의 적용대상이 포함시켜야 한다는 견해도 있고,⁵²⁾ 개인정보보호법의 문리적 해석을 통해 적용가능성 여부를 판단해야 한다는 견해도 있다.⁵³⁾ 그러나 빅데이터에 대해 개인정보보호법을 그대로 적용하는 경우 그 활용에 장애가 있다는 점은 앞서 살펴본 바와 같다. 또한 문리적 해석에 따라 판단하는 경우 입장에 따라 그 해석을 달리하여 혼란을 초래할 수 있다는 점에서 바람직하지 않다.

결국 빅데이터의 활용을 위해서는 엄격하게 보호해야 할 개인정보와 그렇지 않은 개인정보를 분리하여 규정하는 것이 필요하다. 예를 들어 개인의 주민등록번호, 주소 등 신원에 관한 정보는 빅데이터 환경하에서도 엄격하게 보호해야 할 개인정보에 해당한다. 그러나 빅데이터를

51) 진보네트워크와 경제정의실천시민연합 등 시민단체들은 ‘방통위의 빅데이터 가이드라인 심각한 인권침해를 유발할 수 있다’는 의견서와 진정서를 방통위와 개인정보보호위원회에 전달하였다. 경향신문 2013년 12월 30일자 참조.

52) 신영진, 앞의 논문, 97쪽.

53) 최경진 외, 앞의 논문, 79쪽.

특정 애널리틱스를 통해 분석하여 추출한 정보는 이와 구분할 필요가 있다. 예를 들어 앞서 살펴본 아마존닷컴의 경우 개별 고객의 구매내역과 특정제품 구매 후 후속 구매물품에 대한 자료를 분석하여 다음의 애널리틱스에 따라 어떤 제품을 추천할 것인지를 실시간으로 판단하게 된다. 이때 추천하는 상품에 대한 판단은 어떤 애널리틱스를 사용하는가에 따라 달라질 수 있으며, 이를 통해 얻은 정보는 해당기업의 노력에 의해 창출해지는 2차적 정보라고 할 수 있다.

이러한 2차적 정보는 현행 법의 해석 하에서도 개인정보에 포함되지 않는다고 할 수 있다. 현행 법은 ‘해당 정보만으로 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것’을 개인정보에 포함하고 있다. 이때 ‘쉽게’ 결합한다는 것은 각각의 정보 결합의 수단·방법이 합리적으로 이루어질 수 있다는 의미이며, 식별을 위해 불합리할 정도의 시간, 노력, 비용이 투입되어야 한다면 그런 단편적인 정보들은 식별성이 있다고 보기 어렵다.⁵⁴⁾ 따라서 특정한 애널리틱스를 통해 정보를 추출하는 빅데이터 활용은 개인정보에 포함되지 않는 것으로 이해할 수 있다. 그러나 이를 해석에 맡기는 경우 당사자 간 혼란이 야기될 수 있다. 따라서 보호대상 개인정보와 그렇지 않은 개인정보, 예를 들어 공개정보나 2차적 정보를 입법적으로 명확하게 구분하는 것이 필요하다.

3.2.2 정보주체의 동의 방법 구분

개인정보보호법상의 개인정보를 보호대상 개인정보와 공개정보 또는 2차적 정보를 구분하여 규정해야 하는 가장 큰 취지는, 양자 모두 보호가 필요하지만 그 방법을 달리해야 하기 때문이다. 신상정보와 같은 개인정보는 빅데이터 환경하에서도 절대적으로 보호해야 할 정보이며, 따라서 현재와 같이 opt-in 방식으로 정보주체의 동의를 얻는 것이 필요하다.

54) 행정안전부, 앞의 책, 9쪽.

그러나 공개정보나 2차적 정보는 빅데이터 활용을 위해 그 규제를 완화할 필요가 있다. 그렇다고 해서 이를 보호대상에서 완전히 제외하는 경우 개인들의 피해가 우려되며, 결국 개인들이 정보제공 자체를 꺼리는 결과를 가져올 수 있다. 따라서 이를 적절하게 조절할 필요가 있으며, 그 방안으로 생각할 수 있는 것이 opt-out 형태로 동의를 얻는 것이다. 즉, 공개정보를 수집하거나 이용하는 경우 및 2차 정보를 생성하여 이용하는 경우는 사후에 이를 정보주체에 알리고, 정보주체가 원하지 않는 경우에만 이를 활용하지 못하도록 하는 것이다. 기업 입장에서 빅데이터를 자유롭게 활용할 수 있도록 하면서, 정보주체가 원하는 경우 해당 정보의 활용을 제한하여 개인정보도 보호할 방안이 될 수 있다.

3.2.3 ‘잊혀질 권리(right to be forgotten)’의 도입

빅데이터는 기본적으로 개인정보를 바탕으로 형성된다. 대부분의 빅데이터는 개인들에 관한 정보를 바탕으로 새로운 정보를 추출하는 것으로서 개인정보와는 떼어 수 없는 관계이다. 따라서 개인 입장에서는 빅데이터의 활용 필요성에 동의하면서도 자신도 모르게 공개된 정보들이 수집되고 이용될 수 있다는 우려를 하게 된다. 이러한 개인들의 우려는 결국 개인정보 제공을 꺼리는 결과를 가져올 수 있다는 점은 이미 설명하였다.

이러한 문제를 해결하고 빅데이터의 효율적 활용을 위해서는 소위 ‘잊혀질 권리’를 도입할 필요가 있다. ‘잊혀질 권리(right to be forgotten)’란 자신의 정보가 더 이상 적법한 목적들을 위해 필요하지 않을 때, 그것을 지우고 더 이상 처리되지 않도록 할 개인들의 권리를 의미한다. 유럽을 중심으로 논의되기 시작하였으며,⁵⁵⁾ 지난 2012년 1월 EU 집행위

55) 유럽평의회(EC)는 지난해 2010년 4월 11일 유럽평의회 명의의 유럽의회, 평의회, 경제·사회이사회와 종교이사회에 보내는 “유럽연합 내에서의 정보보호에 대한 포괄적 접근(A comprehensive approach on personal data protection in the European Union)”이라는 결정문을 통해 EU 각국이 ‘잊혀질 권리’를 도입할 수

원회가 잊혀질 권리를 명문화하는 내용의 정보보호법 개정안을 확정하였다. 현재 우리나라의 경우도 그 도입 필요성이 계속 제기되고 있는 상황이다.⁵⁶⁾

동 제도는 SNS 등에 자발적으로 올린 글이나 자신에 관한 정보가 삭제되지 않고 계속 기록에 남기는 것이 싫은 경우 개인이 해당 내용을 삭제할 수 있도록 하기 위한 제도이다. 현행법의 경우 개인정보보호법에서 개인정보의 처리정지, 정정 및 삭제 및 파기를 요구할 권리를 보장하고 있다. 그러나 이법의 경우 살아 있는 개인에 관한 정보로서 성명, 주민등록번호 등 식별력 있는 정보로 제한하고 있음에 따라 그 범위가 제한된다. 따라서 법률 개정을 통해 잊혀질 권리를 도입해야 한다는 주장들이 제기되고 있는 것이다.

기업이 빅데이터 활용을 위해 수집하는 개인정보는 구매내역 등과 같은 정보로 민감한 정보는 아니다. 그러나 그 과정에서 자신이 원치 않는 정보를 기업들이 수집할 수 있다는 점에서 불안감을 느낀다.⁵⁷⁾

있는 법적 근거를 마련하였다. EC는 이 결정문에서 그동안 EU각국은 지난 1995년 정보보호지침(1995 Data Protection Directive)이래로 ‘기본권과 개인의 자유의 보호(The protection of fundamental rights and freedoms of individuals)’ 특히 ‘정보보호에 관한 기본권(fundamental right to data protection)’과 아울러 내국 시장에서의 개인정보의 자유로운 유통의 확보(achievement of the internal market - the free flow of personal data)라는 측면을 모두 고려하는 정책을 표방하여 왔으나, 현재에는 개인정보보호와 관련된 주위 환경이 근본적으로 변화함으로써 정보보호의 문제에 대한 인식을 새로이 하여야 할 때가 왔다고 보고 있다. EC, “A comprehensive approach on personal data protection in the European Union”, COM(2010) 609, 2010.4.11.

56) 우리나라의 경우도 많은 문제점이 발생하자 2011.5.5 방송통신위원회가 국가정보화전략위원회에 대해 ‘잊혀질 권리’를 도입할 것이라는 내용을 보고하였으며, 이후 지속적으로 논의가 이어지고 있다.

57) 2012년 2월 미국에서는 한 남자가 자신의 고등학교 딸이 미국 대형유통업체인 타깃(Target)으로부터 유아용품 등 출산용품 할인쿠폰을 광고메일로 받게 되자 매장을 찾아가 항의하는 사건이 발생하였다. 고등학교 딸이 인터넷 검색을 한 이력을 기초로 할인쿠폰을 발행하였으나 해당 부모는 자신의 고등학교 딸이 임신한 사실을 전혀 몰랐던 것이다. 이처럼 부모도 모르는 일을 빅데이터를 통해 알아내고 이를

이러한 불안감을 해소하기 위해서는 본인이 원치 않는 정보를 삭제할 수 있는 권리를 부여하는 것이 필요하다. 즉, 잊혀진 권리를 명문화하여 본인이 원치 않는 정보를 사전에 삭제할 수 있도록 한다면 개인의 불안감을 상당 정도 해소해 줄 수 있을 것이다.

IV. 결 론

빅데이터 시대의 도래는 기업에는 새로운 기회이자 위기라고 할 수 있다. 빅데이터를 잘 활용하는 기업은 경쟁력을 제고할 수 있지만, 이에 도태되는 기업은 그 존립 자체를 위협받을 수 있기 때문이다. 따라서 기업들이 빅데이터를 효율적으로 활용할 수 있도록 제도적인 개선이 이루어질 필요가 있으며, 지나친 규제는 경쟁력 저하의 원인이 될 수 있다는 점에서 신중하게 접근할 필요가 있다.

빅데이터의 활용은 기업운영에 많은 변화를 가져올 것으로 예상된다. 본문에서 살펴본 바와 같이 제품기획, 판매 방식 등에 있어서도 변화가 있을 것이지만, 가장 본질적인 변화는 의사결정구조의 변화를 가져올 수 있다는 점이라고 생각된다. 종래 기업들은 제한된 데이터를 바탕으로 통계적으로 분석하고, 과거 트렌드 분석을 통해 미래를 예측해 보는 시스템을 가지고 있었다. 그러나 이제는 실시간 제공되는 정보를 통해 의사결정을 할 수 있게 됨에 따라 의사결정프로세스에 많은 변화들이 이루어질 것으로 생각된다. 또한 규제 당국의 입장에서도 빅데이터를 통해 기업의 이상 징후를 사전에 인지하고 즉시 대응할 수 있게 됨에 따라 사후적인 조사와 제재중심의 규제에서 벗어나 사전예방적인 규제가 가능할 것으로 생각된다.

마케팅에 활용하게 됨에 따라 개인이 원치 않는 정보의 유출문제가 발생할 가능성이 높아진 것이다. 문상일, 앞의 논문, 24쪽.

그러나 한편으로 개인정보가 침해되는 문제가 발생할 가능성도 배제할 수 없다. 더 큰 문제는 개인정보 침해를 우려하여 개인들이 정보제공 자체를 꺼리는 경우 빅데이터 활용도 불가능해질 수 있다는 점이다. 따라서 빅데이터의 활용과 개인정보보호를 어떻게 조화롭게 가져가느냐가 가장 중요하다고 할 수 있다. 구체적인 방법으로 공개정보나 2차 정보의 수집 및 이용에 관한 절차를 다소 완화할 필요성이 있으며, 또한 잊혀질 권리를 도입하는 것도 고려해야 한다. 그러나 무엇보다도 중요한 것은 빅데이터의 활용이나 개인정보의 보호 중 어느 한쪽에 치우치는 것은 바람직하지 않다는 점이다. 양 측면을 균형 있게 바라볼 수 있어야 실질적으로 빅데이터의 효율적 활용이 가능할 것이기 때문이다.

(투고일 2013년 12월 31일, 심사(의뢰)일 2014년 1월 16일, 게재확정일 2014년 2월 21일)

주제어 : 정보통신기술, 빅데이터, 개인정보, 개인정보보호법, 정보과잉

참고문헌

1. 단행본

최경진 외, 빅데이터 환경에서 개인정보보호 강화를 위한 법·제도적 대책 방안 연구, 개인정보보호위원회 연구용역보고서, 2012.

행정안전부, 개인정보 보호법령 및 지침·고시 해설, 행정안전부, 2011.12.

總務省, 平成24年版 情報通信白書, 2012.

2. 논문

김근태, “빅데이터 분석을 위한 기업의 Big Analytics 환경 변화”, 정보처리학회지 제19권 2호, 한국정보처리학회, 2012, 70-78쪽.

김운석, “개인정보보호 2.0시대의 개인정보보호법 개관”, 법학연구 제22권 제2호, 충남대학교 법학연구소, 2011.12, 4-18쪽.

문상일, “빅데이터 환경에서의 금융소비자 개인신용정보 보호를 위한 제언”, ICT기술의 발전과 경제관련 법제의 쟁점, 한국경제법학회 추계공동학술대회 자료집, 2013, 34-56쪽.

백옥인, “빅데이터의 형성과 전유체제 비판”, 동향과 전망 제87호, 한국사회과학연구소, 2013, 305-330쪽.

성준호, “빅데이터 환경에서 개인정보보호에 관한 법적 검토”, 법학연구 제21권 제2호, 경상대학교 법학연구소, 2013, 307-333쪽.

신영진, “공공분야의 빅데이터 추진과 개인정보보호에 관한 연구: 개인정보 보호법의 주요 내용과 개선과제를 중심으로”, Internet and Information Security 제3권 제3호, 한국인터넷진

- 홍원, 2012, 90-105쪽.
- 안창원/황승구, “빅데이터 기술과 주요이슈”, 정보과학지 제30권 제6호, 한국정보과학회, 2011, 10-17쪽.
- 이광석, “정보재난과 빅데이터 위험 정보사회”, 문화과학, 문화과학사, 2012, 145-165쪽.
- 이규철/원희선, “新기술(빅데이터)등장에 따른 경제적 파급효과 및 법(규제)연구”, 정보와 통신 제29권 11호, 한국통신학회, 2012.10, 48-54쪽.
- 이명진/김우주, “빅데이터를 위한 고급분석 기법과 지원 기술”, Entrue Journal of Information Technology 제11권 제1호, Entrue 정보기술연구소, 2012, 45-56쪽.
- 이창범, “개인정보보호법제 관점에 본 빅데이터의 활용과 보호 방안”, 법학논총 제37권 제1호, 단국대학교 법학연구소, 2013, 509-559쪽.
- 이환수/임동원/조항정, “빅데이터 시대의 개인정보 과잉이 사용자 저항에 미치는 영향”, 지능정보연구 제19권 제1호, 한국지능정보시스템학회, 2013.3, 125-137쪽.
- 장영재, “빅데이터의 활용 -경영을 과학으로-”, 상장협 춘계호, 한국상장회사협의회, 2013, 106-118쪽.
- 최성/우성구, “빅데이터의 정의, 활용 및 동향”, 정보처리학회지 제19권 제2호, 한국정보처리학회, 2012.3, 10-19쪽.
- 奥室 慈, “M2M とビッグデータ時代の到来”, KIIS 144호, 2012, 2-18쪽.
- 風間 博之, “M2M クラウド時代のプラットフォームサービス”, KIIS 144호, 2012, 6-21쪽.
- 城全眞琴, “日本におけるビックデータの現想と課題”, IT Solutions Frontier 29권 3호, 野村総合研究所, 2012, 70-22쪽.
- EC, “A comprehensive approach on personal data protection in the

European Union”, COM, 2010.4.11, 608-633쪽.

John Gantz & David Reinsel, “Extracting Value from Chaos”, *IDC
IVIEW*, 2011, 5-22쪽.

McKinsey Global Institute, “Big data: The next frontier for innovation,
competition, and productivity”, 2011, 1-21쪽.

3. 기타

국가정보화전략위원회, 빅데이터를 활용한 스마트 정부 구현(안), 국
가정보화전략 위원회, 2011.10.26.

[Abstract]

Big Data and Personal Information Protection

Kwan-Hoon Kwak*

There has been a great deal of effort to utilize Big Data around the World and South Korea is not an exception. It is fully expected that using Big Data will become more popular within next 5 years. Due to the explosion of data, coexistence of potential benefit and chaos is highly possible. Thus, many scholars and practitioners have been seeking for a solution to overcome this problem. As commercial companies using Big Data, one of the most common legal issues will be customers' privacy and protection of personal information. More specifically, companies accumulate the consumer's purchase history data which allows the companies to advertise and recommend certain products to those customers. In this case, one of the legal issues that could be raised is whether this type of data or information can be considered as a personal information so that the customers can be protected by a privacy protection act or not. Thus, it is necessary to seek a reasonable resolution which does not yield a strict restriction on utilizing Big Data.

Overall, the current study is to seek a reasonable resolution that use of Big Data and privacy protection can be well-balanced. In fact, private sectors including commercial companies and public sectors are frequently using Big Data. In addition, Big data is closely related to customers and other

* Associate Professor of Department of Police Administration Law at Sunmoon University

interesting parties. Nevertheless, the scope of the current study is limited to utilization of Big Data by commercial companies and privacy or personal information protection.

Key Words : ICT(Information & Communication Technology), big data, personal information, personal information protection law, information overload

