

# KISDI

## Premium Report

### 빅데이터 2.0시대, 주요 이슈와 정책적 시사점

김규남

정보통신정책연구원 부연구위원



## 빅데이터 2.0시대, 주요 이슈와 정책적 시사점

김 규 남 / 정보통신정책연구원 부연구위원

요약문 .....	1
1. 빅데이터의 진화 .....	2
2. 글로벌 빅데이터 정책 동향 .....	5
3. 주요 이슈 .....	9
4. 정책적 시사점 .....	17

## 빅데이터 2.0시대, 주요 이슈와 정책적 시사점

김 규 남

정보통신정책연구원 부연구위원

\*knkim@kisdif.re.kr, 043-531-4315

\*성균관대학교 정보통신공학부 학사

\*서울대학교 기술경영경제정책대학원  
경제학 박사

\*현 정보통신정책연구원

ICT산업연구실

### 요약문

빅데이터 산업은 데이터의 양적 폭발이 일어났던 빅데이터 1.0시대를 지나 빅데이터로부터 실질적 가치를 창출하는 빅데이터 2.0시대로 진입하고 있다. 공급측면에서는 빅데이터 기술이 날이 정교해지고 프로페셔널 서비스가 다양해지는 가운데 빅데이터 전문 기업이 부상하고 있으며, 수요측면에서는 미디어·금융·서비스 산업을 중심으로 대기업의 빅데이터 도입이 빠르게 확산되고 있다.

기업은 빅데이터로부터 실질적인 가치를 창출하고 새로운 비즈니스 기회를 발굴하는데 집중함에 따라 다양한 이슈들이 최근 등장하고 있다. 첫째, 빅데이터의 경제·사회적 파급효과를 분석하고 빅데이터 가치를 정량적으로 수치화하여 평가하기 위한 새로운 개념이 등장하고 있다. 둘째, 데이터가 거래 가능한 자원으로 인식되면서 데이터 거래를 위한 플랫폼이 개발되고 이와 관련된 다양한 형태의 비즈니스 모델이 속속 등장하고 있다. 셋째, 빅데이터 분석 기술이 날이 정교해지고 기업들은 광범위한 데이터로부터 가치 있는 정보를 추출하는데 집중하면서 해킹, 프라이버시 침해 등에 대한 우려도 점차 커지고 있다. 넷째, 빅데이터 전문 인력에 대한 수요가 급증하고 있지만 데이터 과학자, 분석 전문가 등 고급 인력의 공급은 이에 미치지 못하고 있어 인력 수급 불균형 문제 해결이 시급한 과제로 떠오르고 있다.

이러한 이슈들에 대하여 본 보고서는 현재 우리나라의 사회 환경, 시장 환경을 반영한 종합적인 정책적 지원이 필요함을 역설하였다. 구체적으로 빅데이터 산업은 개인정보보호 문제와 맞물려 있으므로 데이터 활용에 대한 사회적 합의 도출이 선제적으로 해결되어야 한다. 또한, 인력수급 불균형 문제 해결을 위해 산업체와 대학 간 실무 중심의 연계체계를 통한 고급 인력 양성의 필요성을 강조하였다. 이와 더불어 빅데이터 육성 정책이 실질적으로 시장에 반영되기 위해서는 민간 기업과의 긴밀한 연계를 통한 시장지향적인 프로젝트를 추진해야 한다. 마지막으로 빅데이터 도입 효과와 정책적 성과에 대한 정확한 측정 방식을 마련하여 정책의 중복성 및 실효성을 점검할 필요가 있다.

## 1. 빅데이터의 진화

### ◆ 빅데이터 2.0시대의 도래

- 데이터의 양적 폭발이 일어났던 빅데이터 1.0시대를 지나 빅데이터로 부터 실질적 가치를 창출하는 빅데이터 2.0시대로 진입
  - 빅데이터의 개념은 데이터의 특성에 기반 한 과거 '3V(양-Volume, 속도-Velocity, 다양성-Variety)'에서 추가적으로 데이터의 가치(Value)에 초점이 맞춰지면서 '4V'로 개념이 확장됨
  - 이는 빅데이터를 하나의 자원으로로서의 가치를 담고 독립적인 산업으로서 근거를 제시하기 위한 것으로 빅데이터의 정의는 데이터를 수집, 저장, 처리, 분석하는 것뿐만 아니라 이로부터 새로운 가치를 창출하는 전 과정을 포괄하는 의미로 확대
- 공급측면에서는 Hadoop 2.2의 등장으로 빅데이터 기술이 정교해지고 빅데이터 벤더들의 프로페셔널 서비스가 다양해지는 가운데 빅데이터 전문 기업이 부상
  - 빅데이터 벤더들은 제품과 서비스의 컨셉을 구체화하고 있으며, 통신, 금융, 바이오 등 특정 산업 분야만을 전문으로 하는 기업들도 증가
  - 빅데이터 HW 벤더들은 SW와 서비스 영역으로 사업을 다각화하고, 리셀러와 같은 非 빅데이터 사업자의 시장 진입으로 벤더 간 경쟁이 치열해지고 있음
  - IBM, HP 등 IT 메가 벤더들이 여전히 빅데이터 시장을 수성하고는 있지만 최근 순수 빅데이터 전문 벤더들이 SW와 서비스 시장을 중심으로 프로페셔널 서비스를 제공하면서 대폭 성장하고 있음
- 수요측면에서는 대기업의 빅데이터 도입이 빠르게 확산되는 가운데 미디어·금융·서비스 산업을 중심으로 빅데이터 수요가 급증
  - 빅데이터 시장 규모는 조사기관마다 다소 차이는 있지만 '15년에는

시장규모가 183~384억 달러에 이를 것으로 전망

※ IDC(2012. 12)에 따르면 빅데이터 기술 및 서비스 시장 규모는 '15년 약 183억 달러에 이르고 연평균 31.7%(CAGR)씩 성장하여 '16년에는 238억 달러에 이를 것으로 전망

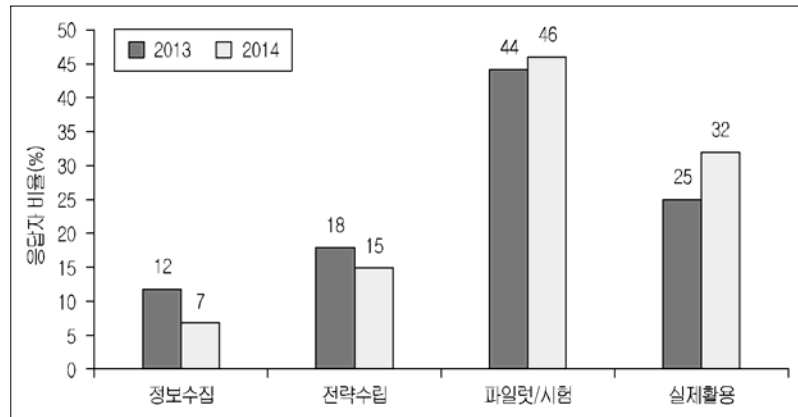
※ Wikibon(2014. 2)은 '15년 빅데이터 시장 규모는 약 384억 달러에 이를 것이며, 연평균 31.7%(CAGR)씩 안정적으로 성장을 하여 '17년에는 최대 501억 달러 규모로 성장할 것으로 전망

- 투자 측면에서는 빅데이터 전문 펀드까지 등장하여 빅데이터 신생 업체에 대한 투자가 크게 증가하고 데이터 분석 솔루션 업체에 대한 관심이 집중되고 있음
- 산업별로는 미디어/커뮤니케이션, 금융, 서비스 산업에서 빅데이터 투자가 적극적으로 이루어지고 있음

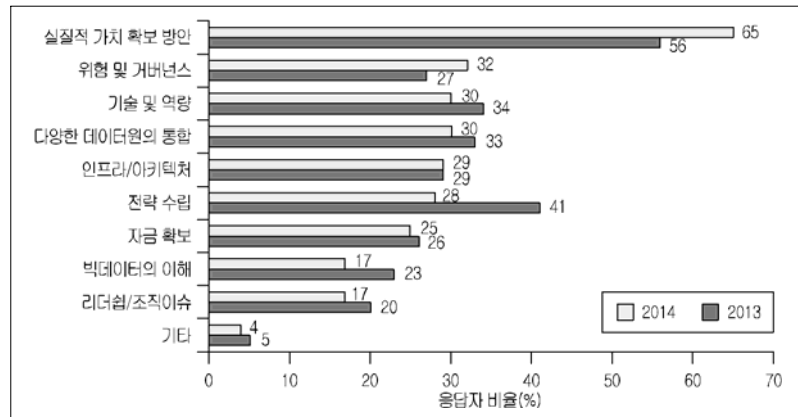
#### ◆ 빅데이터의 실질적 활용 모색

- 빅데이터 도입 기업들은 빅데이터를 단순히 정보수집의 용도로 활용하는 수준을 넘어 파일럿 테스트 등을 통해 빅데이터를 업무에 실질적으로 활용
  - Fortune 1,000 대기업을 대상으로 실시한 빅데이터 활용 현황 조사에서 생산 업무에 빅데이터 이니셔티브를 활용하고 있다는 응답은 67%로 전년도의 32%보다 두 배 이상 증가
  - 기업의 빅데이터 투자는 해마다 점진적으로 증가하고 있으며, 투자 목적도 정보수집과 전략수립보다는 파일럿/시험과 실제 활용을 목적으로 하는 기업이 증가하는 추세임
  - 기업들은 고객경험을 강화하고 업무효율을 개선하기 위해 빅데이터를 도입하고 있으며, 빅데이터의 실질적인 가치를 창출하고 확보하는 방안을 마련하는데 주력하고 있음

(그림 1) 빅데이터 도입 수준과 도입 시 선결과제



(a) 빅데이터 도입 수준



(b) 빅데이터 도입 시 선결과제

자료: Gartner(2014)

**Ford Motors의 사례:** 메이저 자동차 업체인 포드는 빅데이터를 단순한 비용 절감차원이 아닌 변화와 혁신의 수단으로 인식하여 자동차 디자인에서부터 기업의 의사결정에 이르기까지 빅데이터 분석을 활용하고 있음. 일례로 연비를 개선하고 CO<sub>2</sub> 배출을 절감하기 위해 기술 및 시장 데이터를 분석하여 최적의 조합인 경량화와 알루미늄을 채택하였고 이를 제품에 반영하여 철 대신 경량 알루미늄을 사용해 재설계한 트럭(F-150)을 북미 국제 오토쇼에 선보였음.

## 2. 글로벌 빅데이터 정책 동향

- ◆ 빅데이터 정책은 개별 국가의 환경과 발전정도에 따라 세부 전략은 상이하지만 몇 가지 공통적인 중점 추진 방향이 존재함
  - (R&D 및 인력양성) 빅데이터 컴퓨팅 기술 개발을 위한 R&D에 정부가 직접 투자하고 데이터 과학자, 분석전문가, 기술지원 인력 등 전문 인력 양성 추진
    - 미국은 과학기술정책실(OSTP) 주도로 '12년 3월 'Big Data R&D Initiative'를 발표하여 총 2억 달러의 R&D 예산을 빅데이터 분야에 배정하였고, 이후 NITRD(Networking Information Technology R&D)와 함께 후속 프로젝트 진행
    - EU는 유럽 내 과학, 기술, R&D 관련 19개 정부 기관이 참여하여 R&D 및 인력 양성을 추진하고 있으며, 최근 Horizon 2020 Research and Innovation' 프로그램에서는 빅데이터 연구를 위해 25억 유로를 투자
    - 일본은 'Active Japan'(12. 5) 및 '세계 최첨단 IT 국가창조선언'(13. 6)에 빅데이터 정책을 포함하고, '13년도 89.3억 엔의 예산을 배정하여 빅데이터 R&D와 분석 기술 전문가 양성에 투자
  - (산·학·연·관 협력 거버넌스 구축) 산업육성 측면뿐만 아니라 인문-사회-공학의 학제간 융합연구를 통해 빅데이터의 성장 잠재력을 극대화하기 위해 산·학·연·관의 다양한 주체가 협력체계를 구축
    - 미국은 의학에서부터 지리 정보학, 경제학, 언어학 등 다양한 분야에서 빅데이터 프로젝트를 추진하고 있으며, 여기에 하이테크 기업, 제약사, 연구소 등 90여개 이상의 기관이 참여함
    - EU는 빅데이터 프로젝트에 산·학·연·관뿐만 아니라 일반인들도 대규모로 참여하고 있으며, 유럽집행연합(EC)와 빅데이터 가치연합

(Big Data Value Association)과의 민관협력 사업을 추진 중

- ※ 빅데이터 가치연합: 통신사업자 노키아, 네트워크 사업자 오렌지, 전기전자기업 지멘스 및 유럽 내 유수의 연구기관 등 25개 기관이 속한 연합체(한국인터넷진흥원, 2014)
- 싱가포르는 ‘Inforcomm Technology Roadmap’(12. 11)을 수립하여 빅데이터의 장기적 로드맵을 제시하고 여기에 정부 연구기관뿐만 아니라 싱가포르 경영대학, 카네기 멜론 대학 등 대학과 Fujitsu, VMware, EMC 등 다수의 민간 기업이 참여하여 협력하고 있음
- (정부의 공공 데이터 개방) 미국, EU, 영국, 일본 등 주요국은 빅데이터 활용을 촉진하기 위해 정부 주도 하에 공공 데이터를 개방하는 추세임
  - 미국은 ‘Open Data Action Plan’을 통해 거버넌스 혁신을 달성하고 공공 데이터를 개방하여 공공 부문뿐만 아니라 민간 부문의 빅데이터 활용을 활성화
  - EU는 시민의 공공정보 접근권을 보장하기 위해 EU Open Data Portal을 구축하여 공공 부문의 데이터 개방
  - 일본은 ‘오픈 데이터 추진 정책’ 등을 통해 오픈 데이터 전략을 세우고 정부 주도의 공공 데이터 활용을 통한 신 시장 창출에 주력

#### ◆ 우리나라는 선진국과의 수준 격차를 줄이고 빅데이터 산업의 경쟁력을 강화하기 위해 전략적으로 대응해 왔음

- 정부는 ‘빅데이터를 활용한 스마트 정부 구현(안)’(11)을 시작으로 하여 최근 발표된 ‘빅데이터 산업 발전 전략’(13)까지 단계별 추진 과제를 선정하여 빅데이터 인프라 구축과 시장 창출에 주력함
  - ‘빅데이터 서비스 활성화 주요 방안’(12)에서는 신규 서비스 발굴·빅데이터 지원센터 설치·전문 인력 양성 등의 인프라 구축과 개인정보 보호법 개선 등의 제도 개선에 있어 정부 차원의 구체적인 실행 및 지원 계획이 수립되었다는 점이 의의



- ‘빅데이터 산업 발전전략’(13)에서는 서비스뿐만 아니라 관련 산업까지 포괄한 종합적인 산업 육성전략으로 구체적인 목표치가 설정되고 시장 창출 및 확대, 산업 육성 기반 확대, 지속발전 가능한 생태계 조성 등의 세부 전략이 수립되었음

※ 빅데이터 산업 발전전략의 목표: 시장 창출 지원 2배 이상, 핵심기술 개발 7개 분야 선정, 고급인력 양성 5천명 이상, 글로벌 전문 기업 10개 육성 등 4개 항목

- '14년도 국가정보화 시행계획 중 빅데이터 사업은 총 35개 사업에 665억 원의 예산이 배정

※ 35개 사업에는 공간 빅데이터 체계, 특허정보 DB, 빅데이터 플랫폼 등의 인프라 구축 사업과 SW컴퓨팅원천기술개발, 차세대 대용량 처리 기술개발 등의 R&D 사업이 포함됨

- 또한, 정부는 공공 빅데이터 활성화 전략의 일환으로 정부부처가 보유한 공공 데이터를 개방하고 개인 맞춤형 서비스를 제공하는 전자정부 시스템인 ‘정부 3.0’을 추진

〈표 1〉 빅데이터 산업 발전전략

시점	관계부처	빅데이터 산업 정책
2011년	국가정보화전략위원회	‘빅데이터를 활용한 스마트 정부 구현(안)’ 수립·발표
2012년 6월	방송통신위원회	‘빅데이터 서비스 활성화 방안’ 수립·발표
2012년 11월	관계부처 합동	‘스마트 국가 구현을 위한 빅데이터 마스터플랜’ 발표
2013년	미래창조과학부가 설립	빅데이터 정책을 주도
2013년 10월	한국정보화진흥원 (NIA)	빅데이터 인식제고 및 확산을 위해 ‘빅데이터 분석활용센터(현. 빅데이터 전략센터 <sup>1)</sup> )를 개소(2013. 10. 24)

1) 빅데이터 분석활용 센터: 대용량의 데이터를 분석·처리할 수 있는 하드웨어, 소프트웨어 등의 분석

시점	관계부처	빅데이터 산업 정책
2013년 12월	관계부처 합동	빅데이터 산업을 발전시키기 위하여 '빅데이터 산업 발전 전략 (2013. 12. 11)' 발표

자료: 한국데이터베이스진흥원(2014)

플랫폼을 구축, 시범사업 및 공공데이터포털(data.go.kr) 등과 연계하여 다양한 공공 및 민간데이터 세트를 수집·제공함으로써 중소기업의 연구개발 및 사업화 지원, 대학·연구기관의 인력양성에 활용할 빅데이터 활성화 기반 제공을 목적으로 함(한국데이터베이스진흥원(2014))

### 3. 주요 이슈

#### ◆ 빅데이터 가치에 대한 실증적 접근

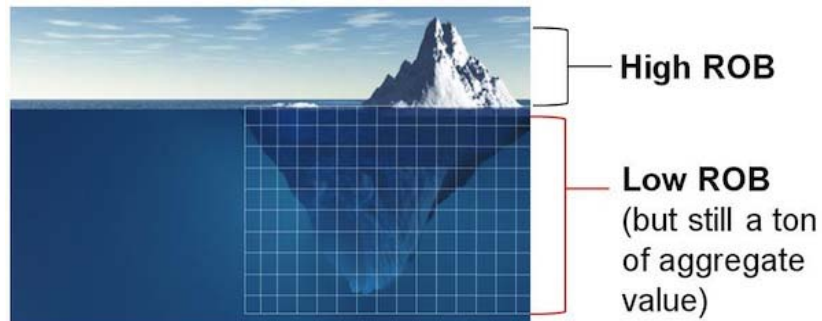
- (거시적) 빅데이터를 경제효과 분석을 위한 수단으로 사용하거나 빅데이터 확산에 따른 경제·사회적 효과에 대한 연구가 활발히 진행 중임
  - 일본(정보통신백서, 2013)은 빅데이터의 축적량과 유통량을 계측하기 위한 프레임워크를 구축하고 실제로 생성, 유통, 축적된 데이터 활용으로 인해 나타나는 경제적 효과를 정량적으로 평가
  - McKinsey는 미국 헬스케어산업, 미국 소매업, 유럽 공공사업, 글로벌 제조업, 글로벌 위치정보데이터 등 5개 부문을 대상으로 산업차원에서 빅데이터의 활용에 의한 경제효과를 연구
  - SAS는 빅데이터 분석을 통해 영국의 조직과 기업에서 창출되는 경제적 가치를 거시 및 산업 차원에서 연구
- (미시적) 빅데이터 가치가 중요해지고 그의 활용 영역이 크게 확대됨에 따라 데이터의 가치를 정량적으로 수치화하여 평가하기 위한 새로운 개념이 등장하고 있음
  - 데이터 수익률(Return on Byte)

$$RoB = \frac{\text{데이터 1Byte당 창출되는 가치}}{\text{데이터 1Byte 저장과 관리에 소요되는 비용}}$$

- RoB가 1보다 크면 데이터에 투입되는 비용 대비 가치가 높아 기업은 데이터를 저장하고 관리하는 반면, RoB가 1보다 작으면 데이터의 가치가 낮아 데이터를 방출하는 경향이 있음
- 데이터가 체계적이지 못하고 노후화될수록 RoB는 낮아질 확률이 높지만, 개별 데이터를 통합하고 가공하여 새로운 정보를 추출해내면 RoB는 높아질 수 있음

- 따라서 빅데이터 2.0시대에서는 가치 있는 정보가 상당히 추출될 수 있는 잠재력을 지닌 원시 데이터를 발굴하고 이를 조직화/체계화 하여 RoB를 최대한 높이는 방향으로 가야함

[그림 2] 빅데이터의 가치 빙산



자료: Awadallah(2013)

#### ◆ 데이터 수익 BM(Business Model)의 다양화

- 데이터가 거래 가능한 자원으로 인식되면서 데이터 거래를 위한 플랫폼이 개발되고 이와 관련된 다양한 형태의 비즈니스 모델이 속속 등장하고 있음
- (데이터 중개) 후지쯔는 '13년 세계 최초로 빅데이터 거래 플랫폼인 '데이터 플라자(Data Plaza)'를 개설하였는데 이는 빅데이터 '중개'라는 사업 모델로 자사 보유 데이터뿐만 아니라 타사 이용자 데이터를 자유롭게 사고 팔수 있는 시스템임<sup>2)</sup>
  - SNS, 통신, 제조, 전자상거래 등과 관련된 기업이 데이터 플라자에 가입할 것으로 기대되며, 후지쯔는 가입 기업 수를 '16년까지 1,000개를 목표로 하고 있음
  - 데이터 플라자에 가입한 기업은 서버에 접속하여 자신이 원하는 데이터를 선택할 수 있고 데이터의 양과 내용에 따라 차별화된 가격

2) 전자뉴스(2013. 1. 30), <http://www.etnews.com/201301300508>

으로 거래할 수 있음

- 이질적인 사업영역에서 기업은 자사에 필요한 데이터를 선택적으로 확보할 수 있어 기업의 경영 효율성 제고 및 수익 개선이 예상되며, 온라인 빅데이터 거래시장도 확대될 전망
- (IT 벤더들의 데이터 수익화) 과거 빅데이터 수요자가 공급자로 전환되는 등 빅데이터 생태계가 역동적으로 진화하는 가운데 이동통신 업체, 웹서비스 업체, IT메가 기업들에 있어 데이터 수익화는 새로운 성장을 위한 핵심 전략으로 부상
  - 데이터를 수집하고 관리하는 데에는 많은 시간과 노력이 투입되지만 일단 양질의 데이터를 확보하면 가공·결합을 통해 다양한 영역에 적용하여 많은 수익을 창출할 수 있어 IT 벤더들의 핵심 사업 모델로 부상
  - Verizon은 '12년 휴대전화 가입자의 인구통계 데이터, 위치정보, 앱 이용 행태와 관련한 데이터 정보를 제공하는 'Precision Market Insights' 서비스를 출시하여 수익 창출<sup>3)</sup>
  - IBM은 수퍼컴퓨터를 이용하여 방대한 의료 데이터를 자동 분석하고 분석 결과를 질병 진단 등 의사의 진료 과정에 참고로 활용할 수 있도록 하는 시스템을 도입하여 헬스케어 서비스 시장에 진출
- (개별소비자의 데이터 수익화) 데이터에 대한 개별소비자의 주권이 확산되면서 소비자가 직접 자신의 데이터를 유통·거래하여 수익화하는 플랫폼이 등장. 미국의 신생 업체 'Datacoup'은 개별소비자 데이터를 축적하여 유통함으로써 수익을 창출하고 이를 다시 소비자에게 배당하는 서비스를 출시하여 주목 받고 있음
  - Datacoup에 가입된 이용자는 SNS를 통해 축적된 개인 프로필, 로그인 기록, 친구목록 등의 개인정보를 신용카드 사용으로 축적된 상점명, 결제액 등의 금융정보와 연계하여 제공. 이때 이용자는 자신의 데이

3) 스트라베이스(2013a)

터를 구매할 업체를 선택할 수 있음

- Datacoup에서는 이러한 정보를 개인비식별화 과정을 거친 후 광고 업체, 전자상거래 업체 등에 판매하여 수익을 창출하고 이를 다시 이용자에게 배분하는데 배당금은 데이터의 가치와 고객 평가에 따라 결정됨
- '14년 2월 기준으로 이미 약 1,500명의 이용자가 등록된 상태이며, 향후 착용형 라이프로그 단말을 통한 데이터나 검색 기록 데이터도 판매할 계획

#### ◆ 프라이버시 침해 점증

- 빅데이터 분석 기술이 나날이 정교해지고 기업들은 광범위한 데이터로부터 가치 있는 정보를 추출하는데 집중하면서 해킹, 프라이버시 침해 등에 대한 우려도 점차 커지고 있음
- 빅데이터로부터 개인식별정보가 추출될 경우 해킹, 프라이버시 침해와 직결되기 때문에 빅데이터의 개인식별화는 빅데이터 산업성장의 가장 큰 위험요소로 간주되고 있음

**개인식별정보**(Personally Identifiable Information: PII)는 단일 개인을 유일하게 식별하거나, 접촉하거나 찾아내기 위해 사용될 수 있는 정보 또는 단일 개인을 유일하게 식별하기 위해 다른 정보원들과 함께 사용될 수 있는 정보를 말함. 개인식별정보가 아닌 정보는 **개인비식별정보**(Non-Personally Identifiable Information: Non-PII)라고 함. 그러나 현대의 뛰어난 재식별(re-identification) 알고리즘은 많은 경우 개인비식별정보만으로도 개인 식별을 가능하게 함(원출처는 Wikipedia(2013. 1), 손상영(2013) 재인용)

- 일부 기업들은 다양한 정보 소스로부터 발생하는 방대한 개인비식별 정보를 정교해진 데이터 마이닝 기술을 통해 개인식별화하고 있으며, 이를 기반으로 표적 마케팅을 하는 것으로 나타나 논란이 되고 있음
- 실제 미국의 일부 헬스케어 기업은 개인 의료 기록을 직접 열람하지 않으면서도 데이터 브로커, 약국, SNS에서 제공된 빅데이터로부터 데이터 마이닝을 거쳐 개인 의료정보를 프로파일링하고 이를 기반으로

임상실험 대상자를 표적 선발한 것으로 드러남(The Wall Street Journal, 2013)

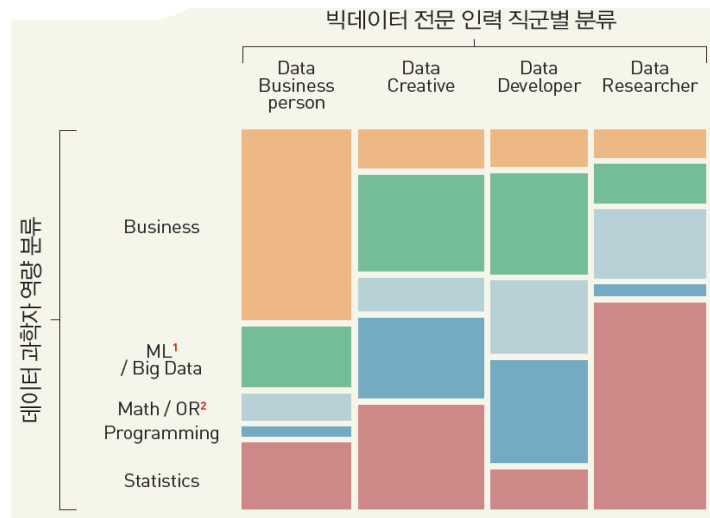
- 미국, EU 등 빅데이터 선진국에서는 프라이버시 침해 방지를 위한 규제 수단을 강화하는 추세이지만 규제 방식, 수행 주체, 제도적 방향성 등에서 차이는 있음
  - 미국은 ‘인터넷 프라이버시 법안(Internet Privacy Act)’(‘10), ‘상거래 프라이버시 권리헌장(Commercial Privacy Bill of Right Act)’(‘11) 등을 통해 인터넷 상에서 발생하는 개인정보를 포괄적으로 보호하고 있지만 근본적으로는 기업의 자율규제를 중시함
  - EU는 ‘데이터 보호 개혁(data protection reform)’(‘12)을 발표하면서 개인 데이터 처리 책임 강화, 데이터 보호의 단일 규정, 잊혀질 권리 등 강력한 규제 수단들을 도입하여 개인정보보호에 대한 국가의 역할을 강화
- 그러나 일각에서는 정보 활용에 관한 지나친 규제는 빅데이터 산업의 혁신과 성장을 저해한다고 주장
  - 미국 소프트웨어 및 정보산업 협회(Software & Information Industry Association, SIIA)는 빅데이터 기술의 다양한 혜택을 주장하면서 정부의 성급한 규제 강화는 오히려 빅데이터 혁신을 지체시킬 수 있다고 우려를 표명함
  - 구글, 아마존, 버라이즌 등 전자상거래 업체들이 지원하는 연구기관인 Technology Policy Institute(TPI)는 사실상 최근 몇 년간 빅데이터의 확산이 명의도용이나 데이터 침해 증가를 유발했다는 증거는 없으며, 새로운 시장에 대한 충분한 고려 없이 마련된 규제는 오히려 혁신에 독이 될 수 있다고 주장

### ◆ 인력 수급 불균형 심화

- 빅데이터 2.0시대에는 데이터 수집·관리도 중요하지만 무엇보다 방대한 데이터 속에서 가치 있는 정보를 찾고 새로운 비즈니스 기회를 발굴하는 데이터 과학자의 중요성이 커지고 있음

- 데이터 과학자는 기술지식, 통계지식, 산업지식을 두루 겸비하고 컴퓨터 분석 및 모델링 기술을 통해 방대한 정보로부터 통찰력 있는 결론을 도출할 수 있는 전문가를 의미
- 마케팅 솔루션 기업인 Sentrana가 전 세계 250명의 빅데이터 전문가를 대상으로 빅데이터 전문 인력 직군별로 데이터 과학자의 역량을 조사한 결과 직군별 요구되는 역량의 비중 차이는 있지만 전반적으로 비즈니스, 기계 학습, 통계 부문에서 일정 수준 이상의 역량이 요구되는 것으로 나타남<sup>4)</sup>

[그림 3] 빅데이터 직군별 데이터 과학자 역량 비중



주: 1) ML-Machine Learning

2) OR-Operation Research

자료: O' REILLY(2013), 스트라베이스(2013b) 재인용

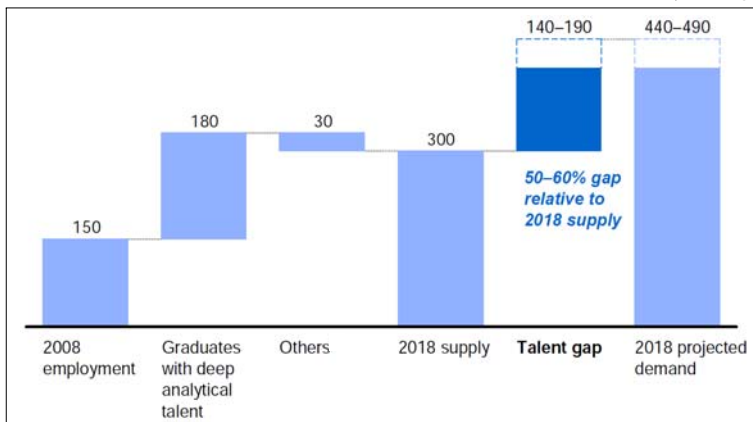
4) 스트라베이스(2013b)



- 빅데이터 기술 발전과 시장 확대로 빅데이터 전문 인력에 대한 수요가 급증하고 있지만 데이터 과학자, 분석 전문가 등 고급 인력의 공급은 이에 미치지 못하고 있어 인력 수급 불균형 문제 해결이 시급한 과제로 떠오르고 있음
  - NewVantage Partners(12)는 금융, 공공, 미디어, 보험 등 38개 기업 및 단체를 대상으로 인력수급과 관련된 설문조사를 실시한 결과 응답자의 60%가 데이터 과학자를 구하기 매우 어렵다고 대답함<sup>5)</sup>
  - 영국 IT관련 비영리단체인 e-skills UK(13)는 영국의 1,000여개의 기업(중소기업 541개, 대기업 500개)과 단체를 상대로 빅데이터 전문가 수요 전망을 조사한 결과 '17년까지 빅데이터 전문가 수요는 69,000명으로 전망되지만 현재 대기업에 재직 중인 빅데이터 전문가는 31,000명에 불과하여 인력 확보가 시급한 것으로 확인됨<sup>6)</sup>
  - McKinsey 보고서(11)에 따르면 미국은 '18년까지 빅데이터 전문 인력에 대한 수요가 약 44~49만 명에 이르지만, 인력 공급은 약 30만 명 수준에 그쳐 약 14~19만 명 정도 인력이 부족할 것으로 전망

[그림 4] 미국의 빅데이터 인력수급 전망

(단위: 천명)



자료: McKinsey Global Institute Analysis(2011)

5) 스트라베이스(2013b)

6) 스트라베이스(2013b)

- 우리나라는 '13년~'17년간 약 52만 개의 빅데이터 일자리가 창출되고, '17년까지 약 11,000명의 빅데이터 전문 인력이 필요할 것으로 예상<sup>7)</sup> 되지만, 현재 빅데이터 전문 인력은 매우 부족하여 인력수급의 불균형이 심각한 상황
  - 국내의 경우 초급 DB 인력은 많지만 중·고급 인력은 부족한 실정으로 삼성경제연구소('12. 5)에 따르면 빅데이터 전문 인력은 100여명에 불과한 것으로 조사됨
  - 국내 금융기업의 빅데이터 조직과 인력 준비 수준도 50.4%에 불과한 것으로 조사되었으며, Cisco('13)가 국내 IT 전문가를 대상으로 한 설문조사에서도 전체 조사자 중 15%가 빅데이터 관련 인력 부족으로 빅데이터 도입이 어렵다고 응답함<sup>8)</sup>
- 이처럼 빅데이터 인력 수급 불균형 문제는 국내뿐만 아니라 선진국들에게도 최우선 당면 과제로 떠오르고 있음

---

7) 한국정보화진흥원(2012, 2013b)

8) 스트라베이스(2013b)

## 4. 정책적 시사점

### ◆ 빅데이터의 실체에 대한 논란에서 벗어나 빅데이터의 실질적 가치를 발굴하고 비즈니스로 연결시켜야 할 시점

- 빅데이터란 무엇인지, 과연 경제적 효과가 있는지 등 빅데이터의 버블 논란은 여전히 지속되고 있지만, 이제 빅데이터는 하나의 산업으로서 자리매김하고 있음
  - 빅데이터 산업은 태동기에서 벗어나 진화 단계에 돌입하면서 기업의 빅데이터 수요는 급증하고 데이터 수익화를 위한 다양한 비즈니스가 등장하고 있음
  - 사물인터넷(IoT)과 클라우드 컴퓨팅의 성장은 데이터 중심 비즈니스(Data-Driven Business)의 확산을 더욱 가속화
  - 빅데이터 도입은 미디어/커뮤니케이션, 금융, 서비스 산업뿐만 아니라 비용 절감 효과를 누릴 수 있는 전자, 소재 산업 등 제조업 전반으로 확산
- 빅데이터 생태계가 본격적으로 구축되어 진화하기 위해서는 사회 환경, 시장 환경을 반영한 종합적인 정책적 지원이 필요함
  - 데이터의 속성은 인적·지역적 환경에 크게 좌우되기 때문에 빅데이터 산업은 지역적 특수성이 강한 산업임. 따라서 빅데이터 생태계를 구축하고 산업을 육성하기 위해서는 우리나라의 사회적 환경, 시장 환경에 적합한 기술공급 및 수요확대 정책을 수립해야 함
  - 이와 관련하여 몇 가지 정책적 제언을 하고자 함

### ◆ 데이터 활용에 대한 사회적 합의 도출

- 빅데이터 산업은 미래 성장 동력으로서 인력양성, 기술개발 등 기술공급 정책도 중요하지만 개인정보보호 문제와 맞물려 있으므로 빅데이터 진흥과 규제 사이 균형점을 찾기 위한 법·제도 확립이 무엇보다 중요함

- 앞서 살펴보았듯이 프라이버시 침해가 점증하는 상황에서 데이터로부터 가치를 창출하고 활용하기 위해서는 사회적 동의에 기반한 제도적 장치가 선제적으로 마련되어야 함
- 프라이버시 침해, 개인 프로파일링 등에 대해서는 강력한 규제를 적용하되 개인 데이터 사용에 대한 새로운 개념 정립을 통해 데이터 활용의 유연성을 강화할 필요가 있음
  - EU와 우리나라는 개인정보 수집·활용에 있어 사전 동의 방식(Opt-In)을 채택하고 있는데, 방대한 데이터가 창출되고 활용되는 상황에서 개인이 일일이 정보보안 정책을 숙지하는 것은 어려울 뿐만 아니라 매사안마다 개인의 동의를 구하는 것은 더 이상 실용적이지 못함
  - 따라서 개인이 단순히 데이터 제공자가 아니라 데이터 통제·처리 주체자로서 개인 데이터 사용에 대한 인식 전환과 더불어 데이터 활용 수준에 대한 사회적 합의를 도출하기 위한 공론의 장이 마련되어야 함
- 이를 토대로 정부는 데이터 관리와 책임 소재 등을 포함한 데이터 수집 및 활용에 대한 명확한 기준을 확립하고 가이드라인을 마련해야 함

〈표 2〉 개인 데이터 사용의 새로운 관점

전통적인 접근	새로운 관점
데이터는 사용자의 인식 하에 능동적으로 수집됨	사물지능통신(M2M) 처리와 수동적인 방식으로 수집된 대부분의 데이터는 개인에게 고지하기 어려움
개인 데이터의 정의가 사전에 결정되었고 이진수로 표시됨	개인 데이터의 정의가 맥락과 사회적 규범에 따라 변함
명시된 사용을 위하여 데이터가 수집되었음	경제적 가치와 혁신은 데이터 집합의 결합과 이후 사용으로부터 발생
사용자는 데이터의 주체임	사용자가 데이터 주체, 데이터 통제자, 데이터 처리자가 될 수 있음
개인이 법적으로 동의하였으나 실질적으로 관여하지는 않음	개인이 데이터의 사용과 가치 창출의 방법에 관여하고 이해함
정책 프레임워크는 개인에게 위험을 최소화하는데 초점이 맞추어져 있음	정책은 보호와 혁신 및 경제적 성장 간의 조화에 초점을 맞춤

자료: WEF&BCG, 한국정보화진흥원(2013a) 재인용

## ◆ 현장 중심의 고급 인력 양성 체계 구축

- 국내 빅데이터 인력 양성은 정부와 민간에서 다양한 방식으로 추진하고 있으나 아직은 초기 단계로 교육 인프라와 실무 능력을 겸비한 고급 인재가 부족한 실정
  - 정부에서는 고급인력 양성을 위한 석사학위과정 개설, 수준별 교육 프로그램 운영, 빅데이터 자격증 제도 도입<sup>9)</sup> 등 다양한 프로그램을 마련하고 있으며, 민간에서는 사내 인력의 재교육 중심으로 빅데이터 인력을 양성하고 있음
  - 그러나 인력 수요에 비해 공급이 턱없이 부족하며, 특히 업계에서 요구하는 데이터에 대한 종합적인 모델링, 분석, 검증이 가능한 데이터 과학자 등 고급인력에 대한 공급이 절실한 상황임
- 기술지식, 통계지식, 산업지식을 겸비한 데이터 과학자를 양성하기 위해서는 산업체와 대학 간 실무 중심의 연계체계를 통한 체계적인 커리큘럼이 마련되어야 함
  - IBM은 조지타운대학, 조지워싱턴대학, 싱가포르국립대학 등 9개 대학과의 제휴를 통해 인력 양성 프로그램을 개설하고 학생들에게 수업 커리큘럼과 더불어 실증적인 빅데이터 역량을 확보할 수 있는 기회를 부여
- 또한, 최근 해외에서는 빅데이터 교육에 온라인 강좌 플랫폼을 활용하고 빅데이터 전문 교육 서비스 업체가 증가하고 있음. 따라서 이를 벤치마킹하여 온라인 강좌, 민간 교육전문기관 활용 등 교육 채널의 다변화를 모색해야 함
  - 미국 MIT는 대학교수, 기술 및 기업 관리자를 대상으로 빅데이터 시장 및 기술 전반에 대한 온라인 강좌 개설<sup>10)</sup>

9) 빅데이터 자격증 제도는 '15년 도입 예정임

10) 스트라베이스(2014)

### ◆ 시장 지향적인 프로젝트 중심의 사업 추진

- 빅데이터 육성 정책이 실질적으로 시장에 반영되기 위해서는 민간 기업과의 긴밀한 연계를 통해 시장의 정책 수요를 정확히 파악하고 시장 환경에 대한 보다 면밀한 조사가 요구됨
  - 시범사업이나 파일럿 프로젝트에는 빅데이터 기술개발자, 분석 전문가, 데이터 과학자뿐만 아니라 비즈니스 전문가도 참여하여 빅데이터의 실질적 활용법을 강구할 필요가 있음
  - 이러한 시장 정보를 기반으로 발전된 빅데이터 기술 개발과 활용을 위한 시장 지향적이고 실용적인 빅데이터 사업을 추진해야 함
- 또한, 데이터 거래가 투명하게 이루어지고 거래로부터 실질적인 효과를 정확하게 파악할 수 있도록 데이터의 가치를 수치화하고 빅데이터 도입 효과를 실증적으로 분석하는 표준적인 방법론을 개발하여 보급해야 함
  - 이를 위해서는 데이터 내용, 형태, 공유범위, 정보보호 등 데이터 활용 원칙이 구체적으로 마련되어야 함

### ◆ 수요·공급 부문에서 정부정책의 실효성 확보

- 미국, EU, 일본 등의 선진국에서는 빅데이터 원천 기술 개발과 더불어 이를 실질적으로 활용할 모델까지 함께 개발하고 있으며, 이를 위해 산·학·연·관 협력 거버넌스를 구축하는데 공을 들이고 있음
  - 특히, 미국은 구체적인 예산 투자 계획을 공표하고 산업 차원에서 적용 가능한 모델을 중심으로 프로젝트를 진행하고 있음
- 우리나라의 경우 다수의 기관에서 여러 차례 빅데이터 산업 육성 정책들이 발표되었기 때문에 이제는 빅데이터 도입 효과와 정책적 성과에 대한 정확한 측정 방식을 마련하여 정책의 중복성 및 실효성을 점검할 필요가 있음

## 참 고 문 헌

- 손상영 (2013), “빅데이터, 온라인 마케팅과 프라이버시 보호”, KISDI 프리미엄리포트 13-01.
- 스트라베이스 (2013a), “이통사와 웹서비스 진영, ‘이용자 데이터’ 수익화 위한 행보 강화, 데이터 주권에 눈 뜬 소비자의 역공 유발”, Trend Watch, 2013. 8. 28.
- \_\_\_\_\_ (2013b), “국내외 기업과 정부의 빅데이터 인력 수급을 둘러싼 현안과 해결책 모색을 위한 행보”, Researcg Prism, 2013. 12. 9.
- \_\_\_\_\_ (2014), “빅데이터 산업 육성 및 인력 수급관련 국내외 정책 추진 현황”, 2014. 10. 28.
- 일본 정보통신백서 (2013), “「스마트ICT」의 전략적 활용으로 어떻게 일본에 활기와 성장을 가져올 것인가?”, 평성25년판.
- 전자뉴스 (2013), “빅데이터 사고파는 장터 생긴다...개인정보보호 문제는?”, <http://www.etnews.com/201301300508>, 2013. 1. 30
- 한국데이터베이스진흥원 (2014), 2014 데이터베이스 백서
- 한국인터넷진흥원 (2013), “오바마 행정부, 빅데이터 R&D 이니셔티브 후속 프로젝트 공개”, Internet&Security Biweekly, 2013년 12월 2회차 (12/12)
- \_\_\_\_\_ (2014), “EU, 3조원 규모의 ‘빅데이터 혁신 기술’ 연구 추진”, Internet&Security Biweekly, 2014년 10월 2회차(10/24)
- 한국정보화진흥원 (2012), “빅데이터 기반의 일자리 창출 전망”, 2012. 12. 20.
- \_\_\_\_\_ (2013a), “빅데이터 시대의 개인 데이터 보호와 활용”, IT&Future Strategy 제8호, 2013. 6. 21.
- \_\_\_\_\_ (2013b), “빅데이터 현황과 전망”, 2013. 11.
- Amr Awadallah (2013). “Big Data’s New Use Cases: Transformation, Active Archive, and Exploration”, Cloudera, (<http://blog.cloudera.com/blog/2013/02/big-datas-new-use-cases-transformation-active-archive>)

-and-exploration/)

EC Press Release, “European Commission and data industry launch €2.5 billion partnership to master big data”, 2014. 10. 13.

e-skills uk (2013). “Big Data Analytics: Adoption and Employment Trends, 2012~2017”.

Gartner (2014). “Survey Analysis: Big Data Investment Grows but Deployments Remain Scarce in 2014”.

IDC (2012). “Worldwide Big Data Technology and Services 2012–2016 Forecast”, 2012. 12.

McKinsey (2011). “Big data: The next frontier for innovation, competition, and productivity”.

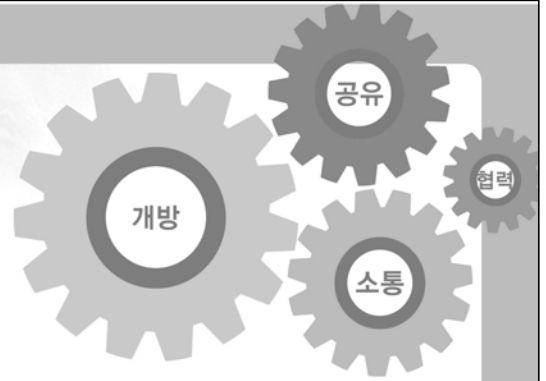
The Wall Street Journal (2013), “정교해진 의료 데이터 마이닝, 사생활 침해 논란”, 2013. 12. 19.

Wikibon (2014). “Big Data Vendor Revenue and Market Forecast, 2013~2017”, 2014. 2. 12.



행복한  
대한민국을 여는

# 정부 3.0



[ 개방 · 공유 · 소통 · 협력 ]

국민의 기대와 희망을 모아 새로운 변화를 시작합니다.  
국민 한 분 한 분을 위해 특별한 내일을 준비합니다.  
개인의 행복이 커질수록 함께 강해지는 새로운 대한민국  
그 희망의 새 시대를 정부3.0이 함께 열어가겠습니다.



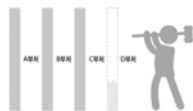
## 국민 모두가 행복해지는 정부3.0



공공정보 공개확대로  
「국민의 알권리」충족



국민의  
정부정책 참여확대



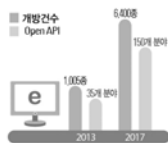
정부 내 칸막이 제거로  
통합적 행정서비스 제공



개인맞춤형 서비스 제공



정보 취약계층  
서비스 접근 제고



공공데이터 민간활용으로  
새로운 일자리창출



정보공유와 디지털협업으로  
더 나은 행정서비스 제공



데이터에 기반한  
과학적 행정구현



창업과 기업활동  
지원 강화



새로운 정보기술을 활용한  
맞춤형 서비스 제공