

공공기관 빅데이터 활용 현황과 과제

신 신 애 | 한국정보화진흥원 빅데이터전략센터

Current status and challenges on the big data of public sector in Korea

Shinae Shin, PhD

Korea Big Data Strategy Center, National Information Society Agency, Seoul, Korea

The ICT (information and communication technology) paradigm shift, including the burgeoning use of mobile, SNS, and smart devices, has resulted in an explosion of data along with lifestyle changes. We have thus arrived at the age of big data. In the meantime, a number of difficulties have arisen in terms of cost or on the technical side with respect to the use of large quantities of data. However, big data has begun to receive attention with the advent of efficient big data technologies such as Hadoop. Big data is recognized as the 21st century digital resource and new driving force. The governments of the United States, Japan, and other countries around the world have already been actively promoting big data in their national policies. Korea is promoting big data actively in the public and private sectors according to National Issues of the Creative Economy and the Government 3.0. In this respect, the Korean government published the “Big Data Industry Development Strategy” in December 2013 and began supporting the use of big data in earnest. This article introduces the status and policy of international and domestic utilization of big data in the public sector. In addition, it describes the challenges in terms of technology, human resources, and data for the vitalization of big data collection and analysis in Korea.

Key Words: Big data; Government 3.0; Creative economy

서론

인터넷, 웹, 모바일, 스마트 기기, 센서 등 정보화의 발전과 ICT (information and communication technology) 패러다임의 변화는 인류의 생활에 큰 변화를 일으키고, 더불어 데이터의 폭증을 가져오게 되었다. 이로 인해 우리는 방대하고 다양한 데이터들이 실시간으로 쏟아지는 데이터 시대에

살게 된 것이다. International Data Corporation은 2011년 전 세계 디지털 정보량은 1.8 제타바이트에 달하는 ‘제타바이트 시대’에 진입하였고, 전 세계의 데이터량은 2년마다 두 배씩 증가하는 것으로 제시하였다.

그동안 대량의 데이터들(빅데이터)을 처리 및 활용하기 위해서는 기술적으로나 비용적인 측면에서 많은 어려움이 있었다. 그러나 최근 빅데이터를 효율적으로 분석, 처리해낼 수 있는 기술들이 등장하면서 실시간으로 빅데이터를 분석하고 관련 데이터를 융복합하여 가치 있는 의미를 찾아낼 수 있게 되면서 빅데이터가 중요한 디지털 자원으로 주목 받기 시작하였다.

빅데이터란 좁은 의미로는 “방대하고 복잡도가 매우 높아 기존의 데이터베이스 도구로는 처리 및 관리가 불가능한 데이터 세트[1]”로 정의된다. Beyer와 Laney [2]는 좀더 넓은

Received: March 31, 2014 Accepted: April 14, 2014

Corresponding author: Shinae Shin
E-mail: sashin@nia.or.kr

© Korean Medical Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

의미에서, 빅데이터를 “큰 볼륨, 빠른 속도 및 높은 다양성(3Vs; volume, velocity, variety)의 특성을 가지고 있어 관리가 어려운 데이터 및 이들 데이터의 축적·처리·분석을 위한 기술과 나아가서는 데이터를 분석해 향상된 의사결정과 유용한 의미 및 통찰, 프로세스 최적화를 가능하게 하는 인재와 조직을 포함하는 포괄적인 개념”으로 정의한다.

그동안 기관들에서는 구조화된 데이터베이스 형태의 정형데이터만을 사용해 왔다면 빅데이터는 웹 문서, 센서데이터 등 반정형데이터와 소셜 데이터, 이미지, 오디오 등 비정형데이터를 모두 포괄하여 분석하고 활용할 수 있게 된 것을 의미하기도 한다. Billings [3]와 Delort [4]는 빅데이터와 기존의 방법들과의 다른 점을 대량의 데이터 세트로부터 규칙(회귀분석, 비선형 관계, 인과효과)을 추론하고 관계와 종속을 밝혀내어 결과와 행동을 예측하는 것으로 구분하고 있다. KT경제경영연구소는 그동안의 데이터 분석은 현상을 이해하기 위해 이론이나 모형에의 (표본)데이터의 부합성을 판단하였다면 빅데이터 분석은 데이터 자체를 통해 패턴과 의미를 도출하고 미래를 예측하는 것으로 기존의 데이터 분석과 빅데이터 분석은 근본적인 차이가 있다고 제시하였다[5].

미국, 일본 등 세계 각국은 이미 정부 차원에서 빅데이터를 국가 정책으로 적극 지원하고 있으며, 우리나라도 2013년부터 시작된 창조경제, 정부 3.0 국정과제 기조에 따라 빅데이터를 신 산업 및 비즈니스 창출, 과학적 행정 및 업무 처리를 위한 기반으로 인식하고 공공과 민간 전반에 국가적으로 적극 추진하고 있다. 이에 본 논문에서는 해외 및 국내의 공공 부문에서의 빅데이터 활용 현황을 소개하고, 우리나라 공공 부문의 빅데이터 활성화를 위한 과제를 제안한다.

빅데이터의 가치 및 해외동향

1. 빅데이터의 등장과 가치

빅데이터는 2010년 세계적 경제지인 Economist에서 빅데이터 특집을 게재하면서 관심이 촉발되었다. Economist는 ‘The data deluge’라는 용어로 빅데이터 시대를 제시하고

데이터는 자본이나 노동력과 거의 동등한 레벨의 경제적 투입 자본으로 비즈니스의 새로운 원자재 역할을 할 것으로 전망하였다[6]. 이후 2011년 Gartner는 빅데이터를 ‘미래 경쟁력을 좌우하는 21세기 원유라 칭하며 데이터가 미래 경쟁우위를 좌우하고 기업은 다가올 ‘데이터 경제 시대’를 이해하고 정보 고립(information silo)을 경계해야 성공할 수 있다고 제시하였다. 또한 Manyika 등[7]은 빅데이터를 경영 혁신 및 사회현안 해결을 위한 핵심 수단으로 의료·행정·소매·제조·개인정보에 빅데이터 적용 시 1% 추가 생산성 향상을 기대할 수 있다고 제시하였다. SAS와 영국의 시장조사기관인 Centre for Economics and Business Research는 빅데이터 도입으로 향후 5년간(2012-2017년) 영국 산업 전체에 약 2,160억 파운드(약 395조 원)의 경제적 효과가 발생할 것으로 전망하였다.

특히, 의료분야에서의 빅데이터 활용은 매우 가치가 높아 연간 3,300억 달러(미국정부 의료예산의 약 8%에 해당하는 규모)의 직간접적인 비용절감 효과를 보일 것으로 전망하기도 한다[7]. 임상분야에서는 의료기관 별 진료방법, 효능, 비용데이터를 분석하여 보다 효과적인 진료방법을 파악하고 환자데이터를 플랫폼화하여 공유하면 치료효과를 제고할 수 있고, 공중보건 영역에선 전국의 의료데이터를 연계하여 전염병 발생과 같은 긴박한 순간에 빠른 의사결정을 가능케 할 것으로 전망된다[8]. 미국 국립암연구소는 암 환자에 대한 빅데이터를 분석해 암 발생의 원인을 밝혔다고 한다. 또한 환자의 상태와 가족력 등을 분석하여 이를 고려한 식단, 운동, 치료가 가능한 이른바 맞춤형 의료의 가능성이 있다고 한다. 이를 통해, 그 동안은 병에 걸리면 치료를 하는 ‘치료’의 개념이 강했다면 빅 데이터는 질병을 차단하기 위한 ‘예방’ 의료를 가능하게 한다.

2. 해외 주요 국가의 빅데이터 정책

세계 각국은 데이터가 새로운 가능성과 혁신의 기회가 될 것임을 인지하고 데이터를 위한 경쟁과 노력을 치열하게 전개하고 있다.

미국 오바마 정부는 대통령 과학기술자문위원회에서 미국 정부기관들이 데이터를 지식으로, 지식을 행동으로 변환하

Toward a leading country using big data

Achieving creative economy, Gov3.0 and citizen happiness through big data

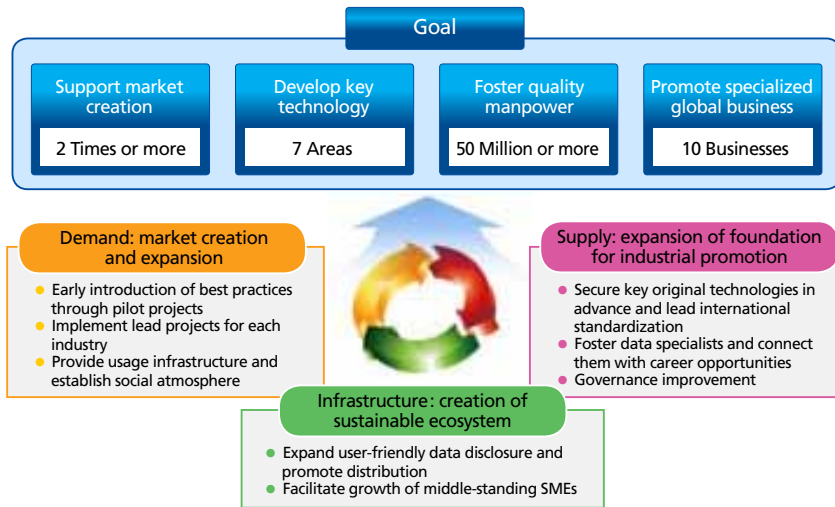


Figure 1. Strategy for big data industry development: vision and major tasks. From Ministry of Science, ICT and Future Planning. Strategy for big data industry development. Gwacheon: Ministry of Science, ICT and Future Planning; 2013 [9]. SME, small and medium sized enterprises.

는 전략에 집중해야 함을 천명하고, 2012년 3월, 'Big Data R&D Initiative'를 발표하였다. 빅데이터 연구개발 이니셔티브는 약 2억 달러 이상을 투입하여 빅데이터 기술개발 및 인력양성을 추진하고 과학기술정책실 주도로 국립과학재단, 국립보건원 등 6개 부처가 협력하여 빅데이터 기술개발 및 국토안보 강화, 질병데이터 분석 등을 수행하며 빅데이터 선진국가로서의 위상을 강화해 나가고 있다. 이중 의료분야와 관련하여 국립보건원은 게놈프로젝트 데이터를 아미존에 무료공개하는 등 신경과학 연구 및 생리학 등 다양한 빅데이터 핵심 기술을 연구개발하고 있다. 또한 휴먼 커넥톰 프로젝트(The NIH Human Connectome Project)를 통해 뇌신경 활동 경로와 기능에 관한 수많은 데이터를 수집하고 이를 도식화하여 전 세계 신경과학연구 데이터에 대한 검색을 도와주는 웹 기반 포털을 운영하고 있으며, 국립보건원과 75개 기업 및 기관들이 협력하는 '1000 유전체 프로젝트'를 통해 유전자 데이터를 공유하고 이를 난치병 치료 및 신약 개발 등에 활용하고 있다.

일본은 빅데이터를 국제 경쟁력 강화를 위한 전략적 자원으로 평가하고 'Active Japan' 보고서를 통해, 빅데이터의 도입으로 10조 엔의 부가가치 창출과 12-15조 엔의 사회적 비용절감이 가능할 것으로 예측하였다. 이에 따라 2012년 5월 '빅데이터 활용 기본전략'을 발표하고 2013년 89.3억 엔

의 예산을 투자하여 데이터 개방, 기반 기술 연구개발 및 표준화, 활용 인재 확보, 사물간 통신 촉진, 규제개선 및 성과평가 방법 마련 등을 추진하고 있다.

또한 영국은 공공기관이 보유하고 있는 공공데이터를 적극 민간에게 개방하여 비즈니스 등에 활용할 수 있도록 지원하는 오픈데이터 정책을 중심으로 데이터 기반 정부 정책을 추진하며 Horizon Scanning Center를 설립하여 기후 변화 예측, 전염병 예방책 마련, 비만 대책 수립 등 영국의 미래를 과학적이고 체계적인 방식으로 연구하고 있다.

국내 빅데이터 정책 및 공공부문 활용 현황

1. 국내 빅데이터 정책

국내 빅데이터 정책은 우선 미래창조과학부가 주도를 하고 있다. 미래창조과학부는 창조경제 실현을 위하여 지난 2013년 12월 Figure 1과 같이 '빅데이터 산업 발전전략'을 발표하고 2017년까지 빅데이터 전문인력 5,000명 양성, 빅데이터 시장 2배 이상 확대를 목표로 수요, 공급, 인프라 측면에서의 정책을 적극 추진하고 있다[9].

수요측면에서는 공공 및 민간에서 관심이 높고 파급효과가 큰 과제를 중심으로 시범사업을 확대한다. 시범사업은 빅데이터 활용이 적합하고 관련 시장을 견인할 수 있는 6대분야(의료·건강, 과학기술, 정보보안, 제조·공정, 소비·거래, 교통·물류)를 정의하여 선도적인 활용 프로젝트를 추진하는 것이다. 또한 한국정보화진흥원의 빅데이터 분석활용센터를 통해 빅데이터 인프라를 개별적으로 직접 구축, 운영하기 어려운 학교 및 중소기업 등이 교육 및 인력양성, 데이터 분석 시 원격에서 활용할 수 있도록 빅데이터 분석 툴과 시스템을 제공한다(Figure 2). 이와 함께 활용도가 높은 공공 및 민간의 테스트 데이터셋(통신, 의료, 센싱 등)을 확보하여 제공함으로써 빅데이터 분석, 활용에 사용할 수 있도록 돕는다.



Figure 2. Korea big data center: vision and strategy.

공급측면에서는 빅데이터 전문인력 양성이 중요한 정책의 축을 이룬다. 빅데이터 선진국인 미국에서는 대학원을 통해 연간 3,000여 명 이상의 데이터 사이언티스트급 전문가를 배출하고 있으나 우리나라는 빅데이터 전문인력 양성이 초 기단계라 할 수 있다. 이에 따라 각 대학/대학원에 빅데이터 관련 학과 또는 과목을 개설토록 유도하고, 빅데이터 전문가 인 데이터 사이언티스트급 인력양성을 지원하기 위하여 관련 커리큘럼도 제시하며, 빅데이터 교육 시 활용할 수 있는 다양한 빅데이터 실습데이터, 모델, 기법 등을 제공하여 빅 데이터 전문인력양성을 지원한다.

인프라측면에서는 공공이 가지고 있는 데이터를 민간에서 활용할 수 있도록 적극 개방을 지원하고 데이터를 유통할 수 있는 마트 등도 마련한다. 정부는 2016년까지 정부가 보유하고 있는 데이터의 60%에 해당하는 총 1만 2,654종의 공공데이터를 개방할 계획이다. 금년도에 2,408종의 데이터를 개방하고 내년에는 2,593종을 추가 개방할 방침이다. 특히 올해는 기상, 국토, 식품의약품, 문화관광 등 산업적 파급효과가 큰 15대 우선 전략분야를 중점 개방한다. 이러한 측면에서 국민건강보험공단은 국민건강정보 데이터베이스를 연구용표본 데이터베이스로 구성하여 대학 및 연구소에서 활용할 수 있도록 개방하고, 건강보험심사평가원도 의료정보지원센터를 개소하고 병원정보 등을 개방하여 민간 및 산업에서 활용할 수 있도록 지원하고 있다.

이와 함께 안전행정부에서도 정부 3.0 기조 아래 빅데이터에 기반한 과학적 행정을 통한 유능한 정부 실현을 위해 빅 데이터 활용 확대 방안을 발표하고[10] 미래창조과학부 등 관련 부처에서 추진하는 빅데이터 활용과제, 관련 법제도 및 활용기반 조성 과제 등을 제시하였다.

2. 국내 빅데이터 현황

정부의 적극적인 빅데이터 정책 추진에 따라 공공부문의 빅데이터 활용 프로젝트는 증가하고 있으며, 전담조직이나

특별조사단을 구성하여 전사적으로 빅데이터를 추진하는 기관들도 증가하고 있다. 특히 국토부, 국방부, 기상청 등은 내부에 빅데이터 활용을 위한 센터 또는 특별조사단을 구성하여 빅데이터 활용을 위한 과제를 적극 추진하기 시작했으며, 국민건강보험공단, 건강보험심사평가원, 한국과학기술정보연구원 등 공공 및 연구소에서도 빅데이터 전문지원조직을 갖추어 나가고 있다.

공공부문에서의 빅데이터는 민원분석, 의료, 교육, 치안, 금융, 환경, 교통 등 다양하게 활용될 수 있다. 민원센터 로그와 소셜 데이터 등에 기반하여 특정 주제에 대한 시민의 목소리를 이해하고 그 추이를 분석하여 정책 의제를 발굴하고 전략을 확보하는 시민 목소리 이해(VOC, voice of customer), 주요 일간지, 소셜 데이터, 민원센터 로그 등을 활용하여 사회이슈의 자동 감지와 연관 주제의 동향 분석 등을 통해 잠재 정책 수요를 발굴하고 지역별 이슈를 도출하여 맞춤형 대국민 서비스 전략 수립, 의료 및 복지 지출 데이터, 민원센터 로그, 소셜 데이터, 서비스기관 홈페이지 및 포털 게시판 등을 활용하여 의료보험 비용을 분석하여 의료복지사업 최적화, 부당청구 방지, 복지정책 입안과 만족도 분석, 지역별 복지 불균형 해소 등에 활용될 수 있다[11]. 또한 교육환경 개선과 교육민원 처리, 합리적 교육예산 집행과 절감, 유행 전염병과 질병에 대한 예측 및 대응과 지역별 분포 분석, 연도별 마크로 분석, 가축 전염병과 환경, 이동 경

로 등에 대한 패턴을 이해할 수 있고 지역별·시간별·이벤트별·유형별 범죄 패턴 분석과 이에 기반한 지역별, 시기별 맞춤형 범죄 예방, 금융 거래 이상 징후 및 조세 회피와 탈세 등의 패턴의 감지와 조기 대응력 확보, 금융 및 조세정책에 대한 시민인식과 지역별, 기간별 동향에 대한 분석 등에도 활용할 수 있다. 그리고 국가 및 도심 센서 네트워크로부터 수집된 환경 데이터와 다양한 보고서 분석을 통해 환경오염과 변화 상황을 모니터링하고, 대응을 할 수 있으며 도로 센서 네트워크, 사건사고 로그, 날씨, 명절 및 스포츠 등의 이벤트 빅데이터에 기반한 교통 흐름 모델링과 예측으로 교통 최적화 시스템 구현 등이 가능하며 센서 데이터, CCTV, 소셜 데이터로부터 도심 내의 문제를 조기 파악하거나 재난을 조기 감지하고 대응하는데 활용할 수 있다.

그동안 개념 중심에 구체적 실체가 부족했던 국내 공공부분의 빅데이터는 2013년 미래창조과학부와 한국정보화진흥원의 공공데이터와 민간데이터를 융합한 6대 시범사업을 통해 실질적인 빅데이터 사례가 제시되기 시작하였다.

서울시는 시내버스 및 정유소 등 교통데이터와 KT의 통신 데이터를 기반으로 한 이동인구 데이터를 융합, 분석하여 최적의 심야버스 노선 정책을 수립하였고, BC카드는 신용카드 거래 데이터, 부동산 데이터, 상가이력 정보 등의 연계 분석을 통해 창업 관련 과거/현황 분석 및 미래예측 정보를 제공하여 창업 의사결정을 지원하기 위한 상권 및 점포평가 시범 서비스를 개발하였으며, 서울대학교와 한국언론재단은 대용량의 기사 데이터베이스에 대해 중요도, 관계도 등 다각도의 고급분석을 적용하여 지능형 뉴스 검색서비스를 개발하였다.

의료분야와 관련하여서 국민건강보험공단은 국민건강주의예보서비스(‘건강IN’ 사이트를 통해 제공)를 개발하였다. 이 서비스는 건강보험진료 현황과 SNS 등의 추세를 분석하여 급성기(전염병) 질병, 계절적·주기적 발생 질병, 다수 및 다빈도 질병 등 건강이상 징후에 대하여 사전에 건강주의예보를 발동하는 것으로 그동안 질병발생 후의 치료중심 서비스에서 건강보험이 예방 중심의 서비스로 전환하는 계기가 될 것으로 기대하고 있다. 또한 한국의약품안전관리원은 ‘의약품안전성조기경보서비스’를 개발하였는데 이 서비스는 유해사례신고 데이터베이스, EMR 정보, SNS·뉴스 등

의 데이터를 분석하여 부작용 등이 의심되는 물질과 약물을 추출하여 조기발견 및 선제적 대응 가능성 여부를 분석하였다. 아산병원은 중환자로부터 수집되는 실시간 심전도·심박수 모니터링 데이터를 분석하여 부정맥 발생 10초 전에 미리 예측이 가능함을 분석하였다. 또한 병원은 병실의 효율적 배분이 경영 및 고객만족에 중요한 요소임에 따라 병원이 보유하고 있는 환자 입원·진료 데이터, 입퇴원 기록, 보험 청구 데이터 등 병원 경영자료를 복합 분석하여 각 과별 병상 배정을 최적화하는 알고리즘을 개발하였다. 계절적 요인, 병원 의료진의 수, 진료 행위의 변화를 모델에 반영하며 이를 토대로 준 실시간 입원 병실의 가동률의 변화 패턴을 파악하여 환자들의 입원 대기시간, 장기재원 일의 감소를 기대하고 있다.

또한 빅데이터 활용이 유망한 과학기술분야에서 빅데이터 활성화를 위해서 미래창조과학부는 2017년까지 사이언스 데이터 맵 구축, 품질관리체계 수립, 분야별 정보 연계, 기술 맵 제시 등을 추진하고 관련 데이터를 공유할 수 있는 오픈 플랫폼 마련을 추진하고 있다.

이외에도 공공부분의 빅데이터 활용 유형은 민원분석, 행정업무 과학화 등으로 나눌 수 있다. 민원이나 국민관심도 분석과 관련해서 문화체육관광부는 그동안은 설문조사 등을 통해 이루어 졌던 국민의식 등을 온라인 뉴스, SNS 등의 빅데이터 분석을 통해 국민의 라이프스타일과 정책에 대한 관심을 조사하였다. 행정업무 과학화와 관련해서는 경상북도는 2013년 미래창조과학부, 한국정보화진흥원과 함께 빅데이터에 기반하여 사과작황정보 분석을 추진하고 있다. 주요 내용은 사과작황에 영향을 미치는 요소들인 농업, 기후, 병충해, 사과관련 웹과 SNS 등 데이터를 수집하여 병충해 및 사과 작황의 패턴 및 추이를 분석하고 관련 요소들의 상관성을 통해 병해충 징후 및 사과 생산 추이, 가격 동향 서비스 등을 제공하고 작황을 예측한다. 이외 다양한 빅데이터 사례는 2013 Korea big data case book [12]을 참조할 수 있다.

3. 국내 빅데이터 활성화를 위한 과제

이제 본격화되기 시작한 국내의 빅데이터가 활성화되고 성공적으로 활용 및 정착되기 위해서는 해결해야 할 중요

한 과제들이 있다. 한국정보화진흥원은 The era of big data that opens a new era [13]를 통해 빅데이터의 성공을 위한 세 가지 요소로 기술, 인력, 자원을 제시하고 있다. 이에 따라 국내 빅데이터 활성화를 위한 과제도 기술, 인력, 자원의 관점에서 정리해볼 수 있다.

첫째, 빅데이터 기술이다. 빅데이터의 분석과 활용은 데이터 수집, 저장 및 관리, 분석, 시각화의 라이프사이클을 가지는데 국내의 빅데이터 기술은 선진 및 글로벌 벤더와 비교해서 2-5년 정도의 기술 격차가 있는 것으로 제시되고 있다 [9]. 그러나 하둡 등 대량의 자료를 분산처리 할 수 있는 오픈소스기술의 등장으로 국내의 빅데이터 기반기술은 빠르게 발전해 가고 있고, 최근에는 한글 처리 및 국내 환경을 고려한 분석, 시각화 기술들도 개발되고 있어 빅데이터 기술을 빠르게 발전될 것으로 기대된다.

둘째, 전문인력양성이다. 빅데이터 속에서 가치를 찾고 활용하기 위해서는 빅데이터 전문인력이 매우 중요하다. 일반적으로 최고 수준의 빅데이터 전문가를 데이터 사이언티스트라고 칭한다. 위키피디아는 데이터 사이언스를 데이터 엔지니어링, 과학적 방법론, 수학, 통계학, 고급 컴퓨팅, 시각화, 해커적 사고방식, 영역별 전문지식을 종합한 학문이라 칭하고, 데이터 사이언티스트를 이를 수행할 수 있는 사람으로 정의하였다. 그러나 아직 국내의 빅데이터 전문가는 매우 부족한 상황이다. 이에, 미래창조과학부와 한국정보화진흥원 등은 대학·대학원과 함께 빅데이터 전문가 양성을 적극 추진하고 이를 지원하기 위하여 대학 빅데이터 교육협의체를 구성하고 빅데이터 인력양성을 위한 커리큘럼 참조모델을 제시하였다(2014년 3월)[14]. 또한, 분야별 전문가 양성을 위하여 교육 시 활용할 수 있도록 다양한 분석 사례와 데이터 셋, 분석 기법과 모델 등도 보급할 계획이다.

셋째, 데이터 자원이다. 빅데이터를 위해서는 풍부하면서도 양질의 데이터가 확보되어야 한다. 다행스러운 것은 우리나라는 정부 3.0 정책과 함께 행정, 과학기술, 기상, 환경, 산업 등 공공부문의 다양한 데이터 개방이 범정부 차원에서 진행되고 있어 풍부한 데이터들을 활용할 수 있는 환경이 마련되고 있고 민간의 데이터도 유통되는 기반도 생성되고 있어 데이터를 자원으로 인식하는 문화가 확산되고 있다.

넷째, 빅데이터 활용 및 서비스 기획력이다. 최근 빅데이터가 주목을 받고 있음에도 불구하고 아직도 많은 공공기관과 민간기업들에서는 빅데이터를 조직에서 어떻게 활용할 수 있을지, 무엇을 해야 할 지 접근법을 찾지 못해 어려워하고 있는 것도 사실이다. 빅데이터 활용은 조직의 필요와 요구에 기반해야 한다. 조직의 핵심 역량과 비즈니스를 식별하고 이에 기반하여 현안과 문제를 도출해야 한다. 그리고 이를 해결하기 위해 빅데이터를 어떻게 활용할 지를 정의하고 필요한 내외부 데이터를 도출하여야 한다. 이때에는 조직의 비즈니스·업무 전문가와 빅데이터 분석 전문가가 함께 팀을 이루어 추진하여야 한다.

다섯째, 개인정보보호이다. 빅데이터에는 개인정보 또는 프라이버시 정보가 포함되어 있을 수 있다. 최근 카드회사의 대규모 개인정보유출 이후 개인정보보호가 빅데이터의 안전한 활용을 위한 전제조건으로 제시되고 있다. 기관·기업에서는 개인정보의 수집은 최소화하고 데이터 수집 단계부터 개인정보의 포함 여부를 식별하고 식별된 개인정보는 암호화 및 비식별화 등을 통하여 안전하게 분석, 활용 하여야 한다. 정부도 빅데이터 산업은 육성하면서도 개인정보는 보호하기 위하여 관련 가이드라인을 마련 중이다. 빅데이터에 대한 국민들의 불안·불신감을 없애기 위하여 빅데이터의 유용성과 안전한 활용 등 대국민 인식제고 및 홍보 등의 노력도 함께 필요하다.

결론

2010년 빅데이터 개념이 본격 등장한 이후 세계 각국은 공공 및 민간에서 빅데이터를 적극 활용하고 있다. 우리나라는 2013년 빅데이터가 정부의 주요 정책으로 등장을 하고 2014년은 빅데이터 활용이 본격화 되는 원년이라 할 수 있다. 2014년 1월 세계경제포럼에서 발표한 미래를 바꿀 10대 신기술 중 하나가 ‘데이터를 활용한 미래예측 모델’이다. 2013년 영국의 시장조사업체 오범(OVUM)이 글로벌 150개 기업을 대상으로 빅데이터 분석 목표를 조사한 결과, ‘의사결정 향상,’ ‘고객서비스 향상,’ ‘보다 정확한 비즈니스

예측' 순으로 조사 되었다. 이는 많은 기업이 경영 프로세스 효율성 제고와 미래 경쟁력 확보를 위해 빅데이터를 고려하고 있다는 것을 의미한다[15]. IBM이 개발한 컴퓨터 왓슨은 2억 페이지의 백과사전, 뉴스, 각종 콘텐츠를 저장하고 복잡한 데이터들의 상관관계를 분석해서 답을 도출한다. 그 결과 2011년 왓슨은 '퀴즈 쇼 제퍼디'에 참가해 상금왕과 다승왕과의 퀴즈대결에서 우승을 차지했고 왓슨은 의료, 금융 등 다양한 분야에서도 빅데이터 처리와 분석기술을 활용하고 있다. IBM의 버지니아 로메티 CEO는 "지금까지 기업들이 의사결정을 내릴 때 주로 CEO나 경영진의 직감에 의존했지만 앞으로는 엄청나게 축적된 데이터의 분석 작업이 직감을 대신하게 될 것"이라고 언급했다.

빅데이터는 데이터가 생성되고 존재하는 모든 분야에서 무궁무진하게 활용 될 수 있다. 우리나라는 부존자원은 부족하지만 해외 어느 나라보다도 정보기술이 빠르게 발전하여 정보기술 사용 능력은 세계적인 수준이며, 수학과 과학 등에 높은 잠재력을 보유하고 있다. 이에 빅데이터는 우리나라에서 경쟁력을 확보할 수 있는 분야로서 공공 및 민간의 창의적 아이디어를 더해 더욱 빠르게 빅데이터를 발전 시킬 수 있을 것으로 기대된다.

찾아보기말: 빅데이터; 정부 3.0; 창조경제

ORCID

Shinae Shin, <http://orcid.org/0000-0002-4305-6456>

REFERENCES

- White T. Hadoop: the definitive guide. Beijing: O'Reilly; 2012.
- Beyer MA, Laney D. The importance of 'big data': a definition. Stamford: Gartner; 2012 [cited 2014 Apr 15]. Available from: <https://www.gartner.com/doc/2057415/importance-big-data-definition>.
- Billings SA. Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. Chichester: Wiley; 2013.
- Delort P. Congres Big Data Paris 2013 [Internet]. Paris: Association Nationale des Directeurs Systemes d'Information; 2013 [cited 2014 Apr 15]. Available from: <http://www.andsi.fr/tag/dsi-big-data>.
- Kim SY. Power to change the world, big data. Seong-nam: KT Economic Management Institute; 2012.
- The data deluge: businesses, governments and society are only starting to tap its vast potential. Economist. 2010. Feb 25.
- Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH. Big data: the next frontier for innovation, competition, and productivity [Internet]. New York: McKinsey & Company; 2011 [cited 2014 Apr 15]. Available from: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.
- Lee SC, Im YS, Ahn M. Big data, secret key to open the future [Internet]. Seongnam: KT Economic Management Institute; 2012 [cited 2014 Apr 15]. Available from: http://www.bicdata.com/bbs/board.php?bo_table=business_archive&wr_id=7.
- Ministry of Science, ICT and Future Planning. Strategy for big data industry development. Gwacheon: Ministry of Science, ICT and Future Planning; 2013.
- Ministry of Security and Public Administration. The plan for expanding the use of big data for smart government. Seoul: Ministry of Security and Public Administration; 2014.
- National Information Society Agency. The guideline for big data analysis and usage version 1.2. Seoul: National Information Society Agency; 2013.
- National Information Society Agency. 2013 Korea big data case book. Seoul: National Information Society Agency; 2014.
- National Information Society Agency. The era of big data that opens a new era. Seoul: National Information Society Agency; 2013.
- National Information Society Agency. Big data curriculum reference model 1.0. Seoul: National Information Society Agency; 2014.
- Jung HM. The power of big data for aware of future, competitive upgrade. The Korea Economic Daily. 2014. Mar 21.

Peer Reviewers' Commentary

본 논문에서는 우리나라의 빅데이터 이용 현황과 전망을 전반적으로 다루었다. 특히 의료, 건강 분야는 향후 빅데이터를 활용하여 가장 획기적으로 변화시킬 수 있는 분야로 강조되었다. 진료 방법의 개선, 치료효과와 비용분석을 통하여 보다 효과적인 진료와 맞춤형 의료가 가능할 것이다. 향후에는 빅데이터 기반의 예방의학이 구체화 될 것이다. 또한 의료보험 비용분석, 부당청구 방지, 복지 불균형 해소에도 빅데이터 분석이 사용될 것이며, 전염병 등 질병에 대한 예측과 대응에도 빅데이터가 활용될 것이다. 본 논문은 우리나라 의료건강 분야의 경쟁력을 확보할 수 있는 새로운 방향과 접근 방법을 제시한 논문으로서 가치가 있다고 판단된다.

[정리: 편집위원회]