

<http://dx.doi.org/10.17703/JCCT.2019.5.2.361>

JCCT 2019-5-45

빅데이터 환경 분석과 글로벌 경쟁력 확보 방안에 대한 연구

A Study on Securing Global Big Data Competitiveness based on its Environment Analysis

문승혁*

Seung Hyeog Moon*

요약 현대 지능정보사회에서 매일 새롭게 만들어지고 유통되는 데이터의 양은 상상을 초월한다. SNS나 인터넷을 통한 데이터로부터 정부나 기업으로부터 창출되는 정보에 이르기까지 다양하다. 이러한 다양한 데이터들이 가공되지 않은 원유와 같이 무한한 가치를 지닌 채 우리 곁에 있다. 축적된 데이터에서 유용한 상관관계를 찾아내고 미래의 불확실성에 대한 예측력을 강화하기 위하여 데이터 마이닝 등을 통한 빅데이터 분석 및 활용은 현대 산업사회에서 모든 영역에 걸쳐 그 중요성이 증대되고 있다. 본 논문에서는 복잡한 현대사회가 생산해 내는 빅데이터의 효율적 관리 및 활용에 대하여 연구한다. 또한 4차 산업혁명시대에 빅데이터를 기반으로 전반적인 산업 경쟁력을 확보하기 위한 전략, 산업 간 시너지 창출 및 비용의 절감과 효과적인 적용방안에 대하여 고찰한다.

주요어 : 빅데이터, 데이터 마이닝, 비정형 데이터, 산업사회, 개인정보

Abstract The amount of data created in the present intelligence information society is beyond imagination. Big data has a great diversity from every information via SNS and internet to the one created by government and enterprises. This various data is close at hand having infinite value as same as crude oil. Big data analysis and utilization by data mining over every areas in the modern industrial society is getting more important for finding useful correlation and strengthening forecasting power against the future uncertainty. Efficient management and utilization of big data produced by complex modern society will be researched in this paper. Also it addresses strategies and methods for securing overall industrial competitiveness, synergy creation among industries, cost reduction and effective application based on big data in the 4th industrial revolution era.

Key words : Big Data, Data Mining, Unstructured Data, Industrial Society, Personal Information

1. 서 론

빅데이터는 21세기의 산업사회에서 없어서는 안 될

귀중한 자원이 되고 있다. 그러나 매일 같이 창출되는 이 디지털 자원은 원유와 마찬가지로 그대로 쓸 수는 없다. 즉 원유를 정제하여 등유, 가솔린, 항공유 등 부

*정회원, 광주대학교 기계·금형공학부
접수일: 2019년 2월 2일, 수정완료일: 2019년 2월 26일
게재확정일: 2019년 3월 10일

Received: February 02, 2019 / Revised: February 26, 2019
Accepted: March 10, 2019

*Corresponding Author: shyoungmoon@gwangju.ac.kr
Dept. of Mechanical and Metallic Mold Engineering,
Gwangju Univ., Korea

가가치 높은 기름을 얻어내듯이 빅데이터도 용도에 맞게 분석과 처리가 이루어져야 국가에 긴요한 자원이 될 수 있다. 유용한 데이터의 확보 및 활용은 4차 산업혁명 시대에 있어 국가 간 경쟁의 근간을 이루는 주요 Enabling Technology인 사물인터넷 (IoT), 인공지능 (AI), 가상물리시스템 (CPS), 클라우드 컴퓨팅 (Cloud Computing) 등 각 영역을 연결시켜주는 핵심 활동이라고 할 수 있다. 따라서 창출된 빅데이터의 수집, 저장, 처리, 분석 및 효과적인 활용을 위한 기술 및 정책 개발은 매우 중요한 과제이다. 빅데이터 구축 및 관련 부문 간의 공유를 통한 산업간 연계성 강화와 글로벌 경쟁력 확보에 대한 연구를 통해 저비용 고효율의 빅데이터 활성화 및 고도화 방안을 제안 한다.

II. 빅데이터의 이해

1. 디지털 데이터로의 진화

정보통신 기기의 눈부신 발달로 디지털 데이터의 생성은 기하급수적으로 늘고 있으며 언제 어디서나 누구라도 손쉽게 만들고 공유한다. 심지어 이러한 데이터와 정보는 인간만이 만들어 내는 것이 아니고 사물인터넷 (IoT)에 의해 생산 공정이나 사물의 움직임에 대한 정보도 인간의 개입 없이 끊임없이 생산되고 있다. 라디오나 TV의 발명은 전파를 통한 음성과 영상 정보의 전송을 가능하게 만들었다. 그러나 한쪽 방향의 일방적 제공이었기 때문에 정보의 양은 방송국의 프로그램 제작 량에 한정되었다. 현재는 인터넷 보급에 따른 P2P 환경 조성과 사용자의 폭증으로 온라인 경제가 확산하면서 상품정보, 구매 데이터가 폭증하고 있으며 블로그나 SNS등을 통한 양방향 소통으로 인해 정보의 생성과 함께 상품 사용자의 성향이나 관계, 구매 행태 등 개인의 취향도 데이터의 형태로 양산되고 있는 상황이다.

2. 빅데이터의 정의

빅데이터는 정보통신 기술의 발달과 사회 현상에 의해 발전되어온 것이므로 모든 사람들이 동의하는 하나의 정의보다는 그 안에 내포된 의미를 통해 알 수 있다. 즉, 성공사례, 속성, 기술적 특징, 추세나 사회, 기업 그리고 사업 프로세스의 영향과 연계하여 설명되어질 수 있다[1]. 빅데이터는 그 크기

가 방대하기 때문에 기존의 단순한 데이터 처리기법을 뛰어넘는 정보의 체계적인 추출, 분석기법을 위한 시스템과 알고리즘을 필요로 한다. 데이터는 지구상의 모든 개인이나 정부, 공공기관, 기업 등과 같은 조직뿐만이 아니고 심지어 사물에서도 생성되며 그 종류도 다양하다. 디지털 환경에서 만들어지는 데이터는 규모가 방대하고, 생성주기가 짧으며 그 형태도 수치 데이터와 같은 정형화된 데이터만이 아니고 음성, 이미지, 영상데이터 등을 포함하는 비정형 데이터로 이루어지는 특성을 가지고 있다. 빅데이터의 속성을 나타내는 것으로 그림 1과 같이 3V (Philip Russom, 2011)를 들 수 있다. 즉, 방대한 데이터의 양을 나타내는 Volume, 데이터의 생성과 전달이 매우 빠른 속도로 진행됨을 나타내는 Velocity 그리고 다양한 형태의 정보를 나타내는 Variety로 특징을 나타낼 수 있다[2].

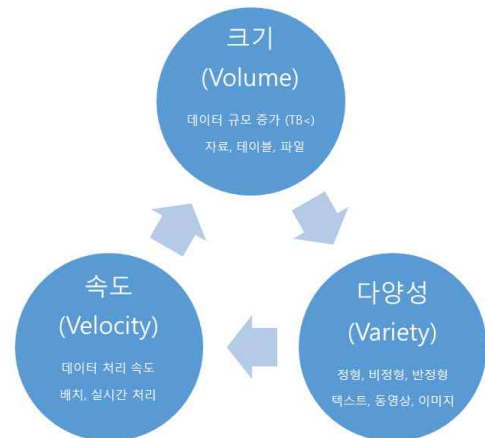


그림 1. 빅데이터의 3대 요소
Figure 1. Three Vs of Big Data

3. 빅데이터 생성속도 및 증가량

빅데이터는 전세계적으로 기하급수적으로 늘고 있어 생성되는 양을 가늠하기 위하여는 그 크기를 측정하는 단위를 이해할 필요가 있다. 글로벌 디지털 데이터 규모는 엑사바이트 이상의 단위로 측정한다. 세계최대의 장서를 보유하고 있는 미국 의회도서관 보유 총 인쇄물 정보량의 100,000배에 해당하는 정보량이 약 1 엑사바이트이다. (Lyman, P., & Varian, H., 2003) 이는 1 기가 바이트의 10억배에 해당하는 수치이며 테라바이트의 백만배, 페타바이트의

1,000배에 해당한다. 그런 엑사바이트의 1,000배에 해당하는 정보량을 1제타 바이트라고 하며 2015년 전 세계의 디지털 데이터 규모는 7.9제타 바이트이며 2020년에는 90제타바이트에 이를 것으로 예상된다.

III. 빅데이터 처리 기술

1. 시스템 구축 방법

빅데이터는 처리해야할 정보가 워낙 방대하여 일반 PC와 같이 처리속도나 저장 용량이 한정되어 있는 하드웨어로는 처리가 불가능하다. 특히 빅데이터에서 의미 있는 데이터의 추출 및 분석은 시의성이 중요한 만큼 빠른 처리속도는 필수적이다. 일반적으로 빅데이터 처리를 위한 하드웨어 구축은 3가지 방법으로 구분되어진다. 첫째, 처리할 빅데이터에 맞게 현재의 컴퓨터를 강력한 컴퓨팅 성능을 발휘할 수 있는 사양으로 시스템을 업그레이드하는 방법이다. 둘째, 여러 대의 컴퓨터를 병렬로 연결하여 데이터를 분산 처리하는 방법이다. 셋째, 클라우드 서비스를 이용, 인터넷을 통해 다수의 컴퓨터를 통해 데이터를 분산 처리하는 방법이다. 세번째 방법은 첫 번째 방법과 두 번째 방법의 단점인 지속적인 시스템 업그레이드나 컴퓨터 숫자의 증가에 따른 공간 확충과 비용에 대한 부담을 경감할 수 있는 장점을 가지고 있다.

2. 빅데이터 처리 기술

1) 하둡 (Hadoop)

빅데이터는 전통적인 데이터 처리 응용 소프트웨어로 다루기에는 너무나 크거나 복잡한 데이터 군집으로, 체계적인 중요 정보의 추출과 분석기법이 필요하다. 앞서 언급한대로 빅데이터는 일반적으로 사이즈가 큰 데이터들을 효율적으로 처리하기 위하여 여러 대의 서버를 통해 데이터를 분산·저장하여 분석·처리를 하는 것이 필요하다. 이와 관련, 구글의 분산 파일 시스템과 병렬 처리 연산 알고리즘인 맵리듀스 (Map Reduce)를 위한 2번째 프로젝트인 아파치 (Apache)를 통해 소개된 하둡은 오픈 소스 기반의 가장 많이 사용되는 솔루션이다. 하둡은 HDFS (Hadoop Distributed File System)을 이용해 빅데이터를 적절히 나누어서 배치하는 마스터 서버 (Name Node)가 할당된 슬레이브 서버 (Data

Node)에 나누어진 데이터를 보내 분산·저장한다[3]. 그림2 에서와 같이 슬레이브 서버에 배치된 파일별 대용량 데이터는 맵리듀스 알고리즘에 의해 처리되어진다. 맵리듀스는 각 파일속의 대용량 데이터로부터 찾고자 하는 유사 데이터를 맵핑하여 찾아내는 과정을 거쳐 의미 있는 데이터를 추출해낸다.

2) 노에스큐엘 (NoSQL)

빅데이터 처리를 위한 전통적인 관계형 데이터베이스 관리시스템 (RDBMS)에서는 데이터를 열 (Row) 과 행 (Column)의 조합을 통해 테이블로 정리하는데, 이 테이블은 독립체를 나타내며 열은 그 독립체의 종류의 하나인 인스턴스 (Instance)를 나타낸다. 각 열은 고유의 키 (Key) 값을 가지며 다른 테이블의 열과 연결이 가능한 구조를 가지고 있다. 이를 위해서는 연결된 열의 고유키를 위한 행을 추가해야 가능하다. 관계는 같은 속성을 가지는 인스턴스의 집합체라고 할 수 있다. 노에스큐엘(NoSQL)은 빅데이터를 유연하게 처리할 수 있는 비관계형 데이터베이스 관리시스템 (DBMS)으로 테이블-컬럼과 같은 스키마 없이, 분산 환경에서 단순 검색 및 추가 작업을 위한 키 값을 최적화하고 데이터 처리 속도가 우수하다. 하둡은 노에스큐엘에 기반을 두고 있는 시스템이다.

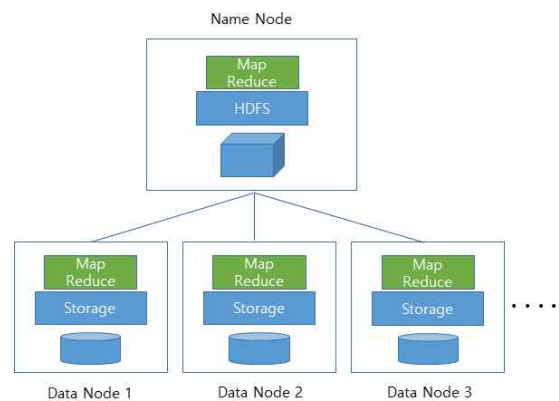


그림 2. 하둡 플랫폼 구조도
Figure 2. Hadoop Platform Structure

IV. 빅데이터 활용과 잠재력

빅데이터의 활용분야는 매우 다양해서 현재 혹은 미래의 고객을 대상으로 제품이나 서비스의 기획, 판매, 마케팅 등에도 적극 활용되고 있다. 예전에는 제품이나 서비스의 마케팅 활동이 불특정 다

수를 겨냥한 매스 (Mass) 마케팅이나, 특정 소집단을 겨냥한 세그먼트 (Segment) 마케팅이 주를 이루었다. 그러나 이는 1:1 마케팅인 CRM (Customer Relationship Management) 마케팅에 비하면 비용대비 효과성 측면에서 뒤떨어진다. CRM은 고객이 원하는 가치를 파악해 이를 충족시키기 위한 제품과 서비스를 지속적으로 제공하여 고객의 만족도를 향상시켜 충성도 (Loyalty)를 높이고 지속적인 관계 유지 (Retention)를 통해 회사의 수익창출에도 기여하는 통합적인 고객관계 관리를 말한다. CRM은 정보통신기술을 활용하여 고객 데이터베이스에서 의사결정에 필요한 데이터를 추출하여 데이터 웨어하우스 (Data Warehouse)에 구축하고 데이터에 숨겨진 상관관계 분석을 위해 데이터 마이닝 (Data Mining)을 시행해 고객 개개인의 특성 분석과 행동 예측 결과를 토대로 해당 고객만을 위한 1:1 마케팅을 가능하게 해준다. 미국의 유료 동영상 스트리밍 서비스 업체인 넷플릭스 (Netflix)는 빅데이터 알고리즘을 이용해 고객이 선호하는 영상물을 추천하기도 하고, 시청자의 취향을 분석하여 영화나 드라마의 연출과 주연 캐스팅에 반영하여 성공을 거두고 있다 [4]. 이렇듯 빅데이터는 고도화된 산업사회에서 필요한 정책이나 사업에 있어서 불확실성의 경감과 빠른 의사결정을 위한 과학적인 솔루션을 제공해 준다.

V. 산업고도화 적용

1. 4차 산업혁명과 빅데이터

4차 산업혁명은 이제 협의의 개념을 뛰어넘어 정치, 경제, 사회, 문화 전반의 미래 지향형 변혁과 인간 개개인의 삶의 질을 향상시켜 주는 광의의 개념으로 진화하고 있다. 즉, 싸고 품질 좋은 제품을 대량으로 만들지만 하면 소비되는 소품종 대량생산에서 다품종 대량생산이나 소량의 주문형 제작이 가능한 매우 유연한 생산체제를 구축, 소비자 지상주의의 사회로 진입하고 있다. 이는 비단 제조업에 국한되는 것이 아니다. 헬스케어 분야에서는 유전체 분석을 통한 미래의 발병 가능 질병을 예측할 수 있는 등 여러 보건의료 분야에서의 국가 비용을 절약해 줄 수 있다[5]. 금융서비스에서는 수익성 높은 투자자문 및 자산관리 서비스를 제공하여 개인의 삶의 질

을 높여줄 수 있다. 4차 산업혁명에 필요한 핵심 기술인 사물인터넷 (IoT), 인공지능 (AI), 가상물리시스템 (CPS), 클라우드 컴퓨팅 (Cloud Computing)에 필요한 자원이 빅데이터인 이유는 4차 산업혁명이 디지털 지식정보에 기반을 둔 산업발전 고도화이며 빅데이터가 이들 기술을 통해 수집, 분석, 가공, 저장, 적용 및 소비되기 때문이다. 국가·단체가 수집한 자료와 사회관계망 서비스에서 배출되는 각종 데이터, 심지어 사물에서 얻어지는 정보는 산업 사회의 고도화를 위한 빅데이터의 원천이 되고 있다.

2. 전망 및 문제점

세계의 빅데이터 시장은 2016년 1,342억 달러에서 2020년 2,100억 달러로 두 자리수 성장이 전망된다. 한국의 빅데이터 시장 규모도 2016년 3,440억 원에서 2020년 9,671억 원으로 연평균 29.5% 성장이 예상되지만 미국(100 기준, 2016년)의 77.3로, 3.3년의 기술격차를 보이고 있고, 일본과도 약 1.9년의 기술 격차를 보이고 있다. (국제무역연구원, 2018) 이는 세계시장에서 한국이 차지하는 비율이 1% 미만으로 풀어야할 과제가 산적해있다고 할 수 있다.

VI. 빅데이터 활성화 과제

1. 빅데이터 확보

수집된 데이터 분석을 통해 국가의 정책 개발이나 기업의 사업전략 수립 등을 위해서는 우선 양질의 충분한 데이터 확보가 필수적이다. 정보통신 기기의 발전 및 보급으로 하루에도 수없이 많은 디지털 데이터가 사회관계망 서비스 (SNS)와 인터넷을 통해 만들어지고 재생산되고 있다. 이는 여러 형태인 숫자, 문자, 영상, 이미지, 음성 등으로 이루어져 있어 이를 용도에 맞게 분류, 가공, 분석하는 데에 많은 인적·물적 비용이 소용된다. 이러한 데이터에는 중요한 정보도 있는 반면 쓸모없는 데이터가 섞여있기 때문에 이를 정제해 내는 것이 중요하다[6]. 아무리 훌륭한 시스템을 도입했다고 하더라도 입수된 데이터의 질이 떨어지면 결과의 신뢰성을 떨어뜨리는 것을 넘어 심각한 정책의 오류나 사업의 실패에 이를 수 있다. 따라서 기업이나 공공기관의 경

우 내부적으로 얻을 수 있는 데이터의 한계를 뛰어 넘기 위하여 외부의 데이터를 확보해야 하는 경우가 많다. 그러나 이 경우에도 비용에 대한 부담이 크며, 예산이 있어도 구하기 어려운 경우가 많다.

2. 빅데이터 분석역량

빅데이터는 이를 수집, 분석 그리고 활용하는 부문이 각각 다른 조직인 경우가 많다. 특히 빅데이터를 사용하는 수준의 기업이나 공공기관은 직능조직이 세분화 되어 있는 경우가 많다. 이 경우 효과적인 데이터의 분석을 위해서는 데이터 수집단계부터 분석 전문가 (Data Scientist)가 데이터의 수집 담당자와의 협력을 통해 용도에 맞는 데이터 확보 계획을 세우도록 해야 한다. 또한 실제로 빅데이터 분석의 결과물을 활용할 부문과도 사전에 활용 목적에 맞는 정교한 분석을 통해 소기의 성과를 달성할 수 있다. 그러나 이러한 분석 역량을 갖춘 인력의 확보가 쉽지 않다. 또한 기업의 경우 유능한 인재가 확보되었다고 해도 분석자에게는 고객과 시장에 대한 이해와 지식이 요구되며 정책이나 전략 그리고 미래에 대한 사업적 통찰력을 지녔을 때 분석된 데이터의 의미 판단 및 시각화를 통해 의사결정을 위한 가치 있는 결과물을 제공할 수 있다.

3. 개인정보 보호

수집된 빅데이터는 다수의 개인정보가 들어 있는 경우가 많다. 이러한 개인정보가 그대로 노출된다거나 다른 목적으로 사용될 경우 개인의 프라이버시 침해는 물론 스팸 메일의 대상이나 심지어 범죄의 표적이 될 수도 있다. 이를 방지하기 위한 개인정보 비식별화 작업이 매우 중요한 이유가 이것이다. 수집된 방대한 데이터를 합법적으로 사용하기 위해서는 사전에 개인별로 개인정보 사용에 대한 동의를 받아야하며 그렇지 않을 경우 어렵게 확보한 데이터의 사용은 불가능하다. 비식별화 작업은 그 자체로 개인을 알아볼 수 있는 데이터는 삭제하거나 누구인지 알아볼 수 없도록 정보를 차단하는 것이다. 현재는 누구인지 식별할 수 없더라도 다른 정보와 결합하면 쉽게 개인을 특정 지을 수 있는 데이터의 경우는 관련된 정보의 변경을 통해 누구인지 알 수 있는 연관성을 없애야한다. 비식별화 시점은

빅데이터를 수집, 분석 및 활용하는 기간 중 개인정보가 식별되거나 다른 데이터와의 결합에 의해 특정 지어지는 경우에는 언제든지 적용되기 때문에 항상 조심스럽게 데이터를 다루어야한다. 이러한 이유로 인해 빅데이터의 사용에 제한이 따르고 있으며 활성화가 늦어지는 주요 요인 중의 하나이다[7].

4. 비용 부담

가치를 담은 데이터는 의미 있는 정보를 지니고 있는 가공이 안 된 재료이다. 수집된 데이터에 내포된 정보도 데이터의 양이 충분해야 비로소 역할을 수행할 수 있으며 다량의 데이터를 확보했다고 해도 이를 처리할 시스템과 전문가를 확보하고 지속적인 운용과 축적된 노하우가 있어야 전락수립이나 의사결정에 필요한 신뢰성 있는 자료를 만들어낼 수 있다. 이러한 물적·인적 비용의 발생으로 인해 중소기업은 물론 웬만한 중견기업조차도 빅데이터를 사업에 적용하거나 확대해 나가는 것에 부담을 가지고 있다. 정부기관의 정책 수립, 통신서비스 사업자의 CRM 마케팅, 금융기관의 고객자산관리 서비스, 의료기관의 임상 연구 분석 등 투자대비 효율성이 뚜렷한 기관이나 조직을 중심으로 활성화되어 있어 빅데이터 활용의 확산을 위해서는 국가의 과감한 정책적 지원이 필요한 분야이다.

VII. 시사점

빅데이터는 정보통신 기술의 발전과 디지털 데이터의 홍수로 인해 우리 곁에 가까이 있고 따라서 많은 사람들이 그 중요성을 인식하고 있는 것이 사실이다. 그러나 현실을 들여다보면 빅데이터가 사업의 실행과 업무의 연관성이 높은 개별 기업이나 공공단체를 중심으로 활용되고 있으며 주로 전략이나 정책적 판단을 위한 자료로서 쓰인다. 세계 주요 국가들은 4차 산업혁명 시대에 제조업에서의 경쟁을 뛰어넘어 광의의 개념에서 국가의 전방위 혁신을 통한 글로벌 경쟁력 제고를 위하여 질주하고 있다. 빅데이터는 이러한 국가적인 노력에 불을 짚어 주는 원동력이라고 할 수 있다. 최근 미국 MIT 대학의 인공지능 단과대학 설립에 대한 세계의 관심이 집중된 가운데, 한국 정부도 인공지능 대학원 설

립을 위한 학교 선정 및 재정지원으로 인공지능 인재 육성에 적극 나서고 있다. 4차 산업혁명 성공에 필요한 인공지능의 발전과 긴밀한 관계에 있는 것이 바로 빅데이터이다. 바둑 기사인 이세돌 9단을 이긴 알파고도 결국 기존의 방대한 기보와 스스로의 강화학습을 통해 만들어낸 빅데이터가 없었다면 인간과의 승리에서 이기지 못했을 것이다[8]. 우리 주변에서 생성되는 많은 정보와 데이터들이 파편화 되어있고 음성, 동영상, 이미지, 메시지 등 여러 형식의 내용으로 이루어진 비정형 데이터인 경우가 많다. 이러한 데이터의 수집, 저장, 처리, 분석 및 활용에는 개별 조직이 감당하기에는 인적·물적 한계가 있어 사용자의 당초 목적에 맞는 신뢰성 있는 결과물을 얻는 것이 어렵다.

VIII. 결론

공공기관이 생성 또는 취득하여 관리하고 있는 데이터를 통합하여 제공하는 창구인 공공데이터포털 (www.data.go.kr)은 정부가 국민들의 편익을 위해 운영하고 있는 것으로 매우 바람직하며 이를 통해 공공 데이터의 접근성이 좋아지고 있다. 그러나 빅데이터는 공공 데이터 외에도 다양한 분야에서 생성되고 있고 필요한 관련 데이터들은 활발히 공유되거나 다른 정보나 데이터와 융합되었을 때 더 큰 부가가치를 만들어 낼 수 있다. 공유의 예를 들면 병원의 질병치료 임상정보는 특수 치료 목적의 의료기기 업체에 제공되면 신제품 개발에 긴요하게 활용될 수 있을 것이다. 융합의 예를 들면 정부의 시내 교통 상황 정보와 자동차 제작사의 차량 운행 정보 및 자동차 보험사의 연령별 사고유형 데이터가 융합되면 한국인에 최적화된 자율주행차 개발 및 안전 운행을 위한 알고리즘이 개발될 수 있을 것이다. 이렇듯 빅데이터의 효과적인 활용에 따라 국가 산업의 발전으로 이어지고 일자리 창출에 도움을 줄 수 있는 기회는 널리 열려있다. 민간 기업의 노력 외에도 정부의 적극적인 지원이나 정책 개발이 필요한 이유가 여기에 있다. 왜냐하면 빅데이터의 수집 및 활용은 국가의 정책과 법규를 기반으로 보다 효율적으로 추진이 가능하고 서로 다른 목

적의 회사나 공공기관의 협력을 이끌어내어 상호 시너지 창출이 용이하며 공공의 이익을 위한 목적이라는 당위성에 부합되기 때문이다. 이와 관련, 정부는 전문 인력 양성과 공공데이터 제공의 확대 외에 다양한 형태의 빅데이터 수급 정책을 통해 개별 업체나 기관의 수요를 충족시켜 줄 수 있도록 해야 한다. 빅데이터 산업진흥과 공공의 이익 추구를 위해서는 빅데이터 사용 환경 조성 및 규제 완화, 빅데이터 활용도에 따른 산업 및 지원 우선순위 분류, 사용자 상호 협력 방안과 글로벌 빅데이터 활용을 위한 지원 등 빅데이터 시장 활성화 및 고도화 방안 마련이 선행되어야 할 것으로 판단된다.

References

- [1] Andrea De Mauro, Marco Greco, Michele Grimaldi, "A Formal Definition of Big Data Based on its Essential Features", *Library Review*, Vol. 65, No. 3, pp.122 - 135, 2016
- [2] Philip Russom, "Big Data Analytics", TDWI Best Practices Report, Fourth Quarter, 2011
- [3] Jaseena K.U., Julie M. David, "Issues, Challenges, and Solutions: Big Data Mining", *NeTCoM, CSIT, GRAPH-HOC, SPTM*, pp. 131 - 140, 2014
- [4] Jacques Bughin, "Big Data, Big Bang?", *Journal of Big Data*, Vol. 3, No. 2, 2016
- [5] Seung-Yeon Hwang, Ji-Hun Park, Ha-Young Youn, Kwang-Jin Kwak, Jeong-Min Park, Jeong-Joon Kim, "Big Data-based Medical Clinical Results Analysis", *The Journal of The Institute of Internet, Broadcasting and Communication (JIIBC)* Vol. 19, No. 1, 2019
- [6]<https://searchbusinessanalytics.techtarget.com/definition/unstructured-data>
- [7] So-Jin Lee, Chae-Eun Jin, Min-Ji Jeon, Jo-Eun Lee, Su-Jeong Kim, Sang-Hyun Lee, "De-identification Policy Comparison and Activation Plan for Big Data Industry", *The Journal of the Convergence on Culture Technology (JCCT)*, Vol. 2, No. 4, 2016
- [8]<https://deepmind.com/research/publications/mastering-game-go-without-human-knowledge/>

※ 이 연구는 2019년도 광주대학교 대학 연구비의 지원을 받아 수행되었음