

빅 데이터의 새로운 고객 가치와 비즈니스 창출을 위한 대응 전략

고준철* · 이해욱* · 정지윤* · 강경식*

*명지대학교 산업경영공학과

Correspondence Strategy for Big Data's New Customer Value and Creation of Business

Joon-Cheol Koh* · Hae-Uk Lee* · Jee-Youn Jeong* · Kyung-Sik Kim*

*Department of Industrial Engineering, Graduate School, University of Myongji

Abstract

Within last 10 years, internet has become a daily activity, and humankind had to face the Data Deluge, a dramatic increase of digital data (Economist 2012). Due to exponential increase in amount of digital data, large scale data has become a big issue and hence the term 'big data' appeared.

There is no official agreement in quantitative and detailed definition of the 'big data', but the meaning is expanding to its value and efficacy. Big data not only has the standardized personal information (internal) like customer information, but also has complex data of external, atypical, social, and real time data.

Big data's technology has the concept that covers wide range technology, including 'data achievement, save/manage, analysis, and application'. To define the connected technology of 'big data', there are Big Table, Cassandra, Hadoop, MapReduce, Hbase, and NoSQL, and for the sub-techniques, Text Mining, Opinion Mining, Social Network Analysis, Cluster Analysis are gaining attention.

The three features that 'big data' needs to have is about creating large amounts of individual elements (high-resolution) to variety of high-frequency data. Big data has three defining features of volume, variety, and velocity, which is called the '3V'. There is increase in complexity as the 4th feature, and as all 4features are satisfied, it becomes more suitable to a 'big data'.

In this study, we have looked at various reasons why companies need to impose 'big data', ways of application, and advanced cases of domestic and foreign applications. To correspond effectively to 'big data' revolution, paradigm shift in areas of data production, distribution, and consumption is needed, and insight of unfolding and preparing future business by considering the unpredictable market of technology, industry environment, and flow of social demand is desperately needed.

Keywords : Big data, Data Deluge, 3V, Hadoop, MapReduce, Social Network

† 교신저자: 고준철, 경기도 안양시 만안구 안양6동 호정타워 913호

M · P: 010-6363-4137, E-mail: tobe0515@empal.com

2012년 10월 20일 접수; 2012년 12월 5일 수정본 접수; 2012년 12월 7일 게재 확정

1. 서론

지식정보화 시대를 넘어 21세기를 감성의 시대라고 일컬으며 정보통신기술의 발전과 함께 스마트 시대라고도 일컫는다. 정보기술(IT: Information Technology)의 일상화가 이루어지는 스마트시대에는 소셜, 사물, 라이프 로그(개인들의 일상생활의 기록) 데이터 등이 결합되며 '빅 데이터'의 영향력이 증대되고 있다. 정보통신 기술은 가상현실, 인공지능, 프로슈머, 바이오 경제 등 사회 변화와 새로운 가능성의 기반이며 드림사이언티, 하이컨셉, 하이터치 등 가치 창출의 근간을 이루고 있다.

실시간 연결과 소통의 '스마트 혁명'은 데이터 폭증을 발생시켰고 기존의 데이터 저장, 관리, 분석기법은 한계와 도전에 직면해 있으며, 데이터는 정보사회를 움직이는 핵심 연료인 만큼, 빅 데이터로의 환경변화는 정보사회의 패러다임을 견인할 정도의 큰 힘을 발휘하고 있다.

인터넷이 일상화된 최근 10년 사이 인류는 디지털 데이터가 폭증하는 데이터 홍수(Data Deluge) 현상에 직면(Economist 2010)하게 되었고, 디지털 정보량의 기하급수적인 증가에 따라 대규모 데이터가 중대 이슈로 부각하며 '빅 데이터'라는 용어가 등장하게 되었다.

'빅 데이터'란 단순한 크기나 양의 측면보다는 기존의 방식으로 수집, 관리, 분석, 시각화, 활용 등이 어려운 새로운 특성의 방대한 정형, 비정형 데이터 세트로 보통 수십에서 수천 테라바이트(TB, Terabyte)에서 수 페타바이트(PT, Petabyte) 정도의 거대한 크기를 생성-유통-소비(이용)가 몇 초에서 몇 시간 단위로 얼마나 기존의 방식으로 관리와 분석이 매우 어려운 데이터의 집합을 의미한다.(빅 데이터 경영을 바꾸다. 한유근, 채승범, 2012, 8, 삼성경제연구소)

빅 데이터는 규모, 다양성, 복잡성, 속도의 증가 특성을 가지고 있으며, 4개의 요소가 충족되었을 때 빅 데이터에 적합하다.

미래사회가 지향하는 국가 혁신체계는 지능형 인프라, 융합 시스템, 스마트 서비스를 기반으로 발전하고, 개방과 협력, 통합적 시스템으로 진화하고 있다. 특히 IT 패러다임의 변화는 데이터의 양적 팽창을 가져왔으며, 데이터 범람이 새로운 기회와 편익을 창출해 내기 시작했다. 또한 빅 데이터는 잠재적 가치와 잠재적 위험이 공존하며, 사회, 경제적으로 성과를 좌우하는 핵심원천이 될 것으로 평가하고 있다.

국내 뿐만 아니라 세계 각국의 정부와 기업들은 빅 데이터가 향후 기업의 성패를 가늠할 새로운 경제적

가치의 원천이 될 것으로 기대하고 있다. 빅 데이터에서 유용한 정보를 찾아내고 잠재된 정보를 활용할 수 있는 기업들이 경쟁에서 시장을 선도할 수 있는 기회를 확보하게 된다. 한편, '정보의 바다'에서 '데이터 홍수의 시대'로 변화하는 시대에 빅 데이터에 대처하지 않으면 상당한 위험에 노출될 것으로 우려하고 있다.

사회 발전의 속도가 빨라지고 위험요인과 복잡성이 증가할수록 시스템적으로 신속하게 환경 변화를 감지할 필요성이 증가하고, 빅 데이터의 분석을 통해 통찰력, 대응력, 경쟁력, 창조력을 향상시키며 기업이나 국가가 지속적으로 성장 발전을 이룰 수 있는 전략 수립이 필요한 때이다. 급변하는 글로벌 환경 및 미래사회에 선제적으로 대응하기 위해서는 대규모 데이터 분석 기반의 국가 미래 전략 수립 체계 마련이 선결 과제이다. 또한, 국가나 기업이 다가올 '데이터 경쟁 시대'를 이해하고 대비하는 '연결과 협력'체계 마련과 빅 데이터의 핵심 역량인 '창의적 인력' 확보와 내부역량 강화에 노력을 경주해야 하며, 데이터의 안전한 사용을 보장하고, 데이터의 신뢰성을 높이기 위한 법적, 사회적 측면의 환경 조성이 필수적이다.

따라서 본 연구에서는 빅 데이터 비즈니스란 무엇이고, 빅 데이터 활용가치 및 국내외 적용 사례를 살펴보고 새로운 고객가치와 비즈니스 창출을 위해 국가나 기업이 준비하고 대응해야 할 전략에 대해 정의하고 빅 데이터 확산에 따른 도전과 기회를 제공하고자 한다.

2. 이론적 고찰

2.1 빅 데이터 비즈니스 정의

빅 데이터의 정의는 데이터의 규모와 기술 측면에서 출발했으나, 사실 빅 데이터에 대해서 구체적이고 정량적인 정의가 합의된 바는 없으며, 빅 데이터의 가치와 활용효과 측면으로 의미가 확대되는 추세이다. 빅 데이터는 고객정보와 같이 정형화된 자산정보(내부) 뿐만 아니라 외부 데이터, 비정형, 소셜, 실시간 데이터 등이 복합적으로 구성되어 있다. 정보통신 분야의 커다란 조류의 하나로 빅 데이터 비즈니스의 시작이 예상되며 이에 따라 빅 데이터를 정의하면 '기존의 관리 및 분석 체계로는 감당할 수 없을 정도의 거대한 데이터의 집합을 지칭(SERI, 2010년)하며, 대규모 데이터와 관계된 기술 및 도구(수집, 저장, 검색, 공유, 분석, 시각화 등)도 빅 데이터의 범주에 포함하고 있다. 또한 스키프 료스케는 그의 저서 '빅 데이터 비즈니스'에서 '사업에 도움이 되는 인사이트를 도출하기 위해 고해상, 고빈도로

생성되는 다양한 데이터(2012, 더 숲, p23)로 정의하고 있다. 또한 데이터량 기준에 대한 산업분야에 따라 상대적으로 현재 기준에서는 몇 십 테라바이트에서 수 페타 바이트까지가 그 범위로 정하고 있다.(KT 중앙연구소 Intelligent Knowledge Service, 조성우)

빅 데이터의 데이터 종류는 정형화 정도에 따라 다음과 같이 분류한다.

<표 2.1> Types of Big Data

정의	설명
정형	고정된 필드에 저장된 데이터. 관계형 데이터베이스 및 스프레드시트 등을 예로 들 수 있다.
반정형	고정된 필드에 저장되어 있지는 않지만, 메타데이터나 스키마 등을 포함하는 데이터. XML이나 HTML 텍스트 등을 예로 들 수 있다.
비정형	고정된 필드에 저장되어 있지 않은 데이터. 텍스트 분석이 가능한 텍스트 문서 및 이미지/동영상/음성 데이터 등을 예로 들 수 있다.

2.2 빅 데이터 기술의 이해

2.2.1 차세대 이슈, ‘빅 데이터’

정보통신기술(ICT) 전체를 흐르는 데이터의 변화와 발전은 IT와 비즈니스 영역에 새로운 공급체제와 수요를 창출하는 원동력으로 작용하고 있는데, 정보통신기술 인프라의 성숙에 따른 신규 투자처에 대한 니즈 증가로 새로운 공급체제와 시장을 창출하는 빅 데이터에 대한 관심이 증가하고 있다는 것과 기업은 그간 축적된 자산 데이터와 소셜 네트워크에 기반 한 대규모 외부 데이터의 활용으로 혁신과 기회포착을 기대하고 있다. 또한, 인적 관계 데이터, 위치, 형태, 인식 등과 같이 사람과 사회현상을 이해할 수 있는 비정형의 대규모 데이터가 증가하고, 정보가 중요한 자산인 지식정보 시대에는 대규모 데이터에서 의미를 찾고 정보와 지식을 만들어 내는 능력이 바로 경쟁력이다.

2.2.2 빅 데이터 기술의 분류

빅 데이터 기술은 ‘데이터의 획득, 저장, 관리, 분석, 활용’ 전반에 걸친 광범위한 기술을 포괄하는 개념이다. 빅 데이터 기술은 크게 기존 기술과 새로운 기술로 구분하고, 전자는 빅 데이터라는 현상이 주목을 받기 전부터 존재해왔던 데이터 관리 언어 SQL(Structured Query Language : 관계형(구조화) 질의어)과 관련된

기술들이다. 반면 후자는 빅 데이터가 등장하는데 결정적인 역할을 한 신규 기술들로 좀 더 유연하지만 복잡한 데이터 관리 언어 NoSQL(SQL을 사용하지 않는 언어:비 관계형 질의어)을 기반으로 한다. 이는 좁은 의미의 빅 데이터 개념을 반영하는 기술로 비구조적 데이터 형태를 획득, 조직화, 분석하기 위한 분산파일 시스템, 맵리듀스(MapReduce)솔루션(Hadoop), NoSQL 등을 의미한다.

빅 데이터와 연계된 기술들을 정의하면 다음과 같다.

(1) 빅 테이블(Big Table)

구글 파일 시스템 상에 구축된 상용 분산 데이터베이스 시스템으로 H베이스(Hbase)에 영향을 줌

(2) 카산드라(Cassandra)

분산 시스템에서 방대한 분량의 데이터를 처리할 수 있도록 디자인 된 오픈소스데이터베이스 관리시스템

(3) 하둡(Hadoop)

분산 시스템상에서 대용량 데이터 처리 분석을 지원하는 오픈소스 소프트웨어 프레임워크로 구글이 개발한 맵리듀스를 오픈소스로 구현한 결과물로 현재 정형/비정형 데이터 분석에 가장 선호되는 솔루션(HDFS: 하둡분산파일시스템), Hbase, MapReduce등이 포함)

(4) 맵리듀스(MapReduce)

분산 시스템상에서 대용량 데이터 세트를 처리하기 위해서 구글이 제안한 소프트웨어 프레임워크로 하둡에도 구현되어 있음

(5) H베이스(Hbase)

구글의 ‘빅 테이블’을 참고로 개발된 오픈 소스 분산 비 관계형 데이터베이스다.

(6) NoSQL(비 관계형 데이터베이스)

비 관계형 데이터베이스는 데이터를 테이블에 저장하지 않는 데이터베이스이며 관계형 데이터베이스와는 대조적인 개념으로 NoSQL의 대표적인 솔루션으로 Cassandra, Hbase, MongoDB 등이 있다.

2.2.3 빅 데이터 분석 기법

정형/비정형 데이터가 예측할 수 없을 정도로 증가하고 있으며, 특히 비정형 데이터의 증가 속도가 급속히 빨라지고 있다. 이에 따라 빅 데이터를 처리하기 위해서는 기존 데이터와 빅 데이터 처리에 대한 차이점을 살펴볼 필요가 있다. 첫째, 빅 데이터는 기존 데이터에 비해 빠른 의사결정이 덜 요구된다. 기존의 데이터 처리에서 요구되는 즉각적인 처리속도와는 달리, 즉각적인 의사결정이 상대적으로 덜 요구된다. 둘째, 다양한 데이터 소스, 복잡한 로직처리, 대용량 데이터 처리로

인해 처리 복잡도가 매우 높다. 셋째, 처리할 데이터 양이 방대하다. 넷째, 비정형 데이터의 비중이 높다. 다섯째, 대용량 및 복잡한 처리를 특징으로 하고 있어, 동시에 처리(Throughput)가 필요한 데이터양은 낮다. 최근 소셜 미디어 등 비정형 데이터의 증가로 인해 분석기법들 중에서 다음과 같은 분석기법이 주목을 받고 있다.

(1) Text Mining

비/반정형 텍스트 데이터에서 자연어 처리 기술에 기반하여 유용한 정보를 추출, 가공하는 것을 목적으로 방대한 텍스트 문치에서 의미 있는 정보를 추출해 내고, 다른 정보와의 연계성을 파악하며, 단순한 정보검색 이상의 결과를 얻어 낼 수 있다.

(2) Opinion Mining

소셜 미디어 등의 정형/비정형 텍스트의 긍정, 부정, 중립의 선호도를 판별하는 기술이다.

(3) Social Network Analysis

소셜 네트워크 연결 구조 및 연결 강도 등을 바탕으로 사용자의 명성 및 영향력을 측정하여 소셜 네트워크상에서 입소문의 중심이나 허브 역할을 하는 사용자를 찾는데 주로 활용된다.

(4) Cluster Analysis

군집분석은 비슷한 특성을 가진 개체들을 합쳐가면서 최종적으로 유사 특성의 군을 발굴하는데 사용된다.

2.3 빅 데이터의 세 가지 특징

개별적이고 즉각적이며 다면적인 검토를 거친 부가 가치를 제공하고 싶어 하는 요구를 만족시키는 인사이트를 ‘사업에 도움이 되는 인사이트’라고 정의할 때 빅 데이터가 지녀야 할 3가지 특징을 살펴보면 수많은 개별 요소 각각에 관한 것(고해상)을 고빈도로 생성(고빈도 생성)한 다양한(다양성) 데이터를 일컫는다.

2.3.1 고해상

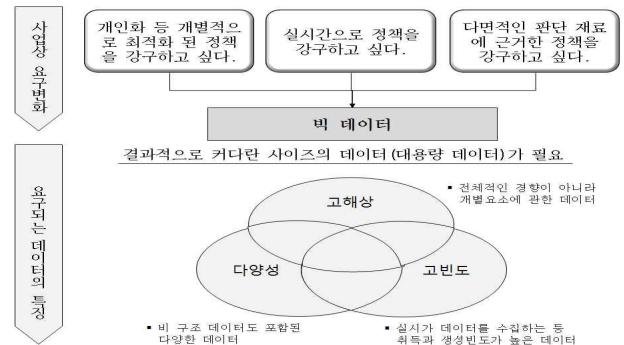
기존에는 한데 묶어서 다루어왔던 현상을 각각의 요소로 분해하여 파악하고 대응할 수 있는 데이터여야 한다.

2.3.2 고빈도 생성

취득이나 생성 혹은 처리 대상이 되는 데이터의 사이즈가 크지는 않더라도 매우 높은 빈도로 생성된다면 이것은 기존에는 불가능했던 실시간 정책의 실시를 가능하게 한다.

2.3.3 다양성

정형적인 수치 데이터, 텍스트 데이터에 그치지 않고 웹 서비스를 이용한 사용자의 기록, 방법 카메라 영상, 위치정보, 각종 센서에서 수집된 데이터 등 다양한 종류의 데이터를 연계하여 활용하는 것도 요구된다.



<그림 1> Three characteristics of Big Data

2.4 빅 데이터의 4가지 구성요소(특성)

데이터 분석을 통하여 가치 있는 정보를 추출하는 비즈니스 방향은 빅 데이터와 함께 출현한 새로운 현상이 아니다. 그럼에도 불구하고 빅 데이터 현상이 다루는 데이터는 기존의 데이터와 다른 특성들을 가지고 있다. 빅 데이터는 규모(Volume), 다양성(Variety), 속도(Velocity)의 이른바 ‘3V’와 복잡성(Complexity)의 증가 특성을 가지고 있으며 4개의 요소가 충족될수록 빅 데이터에 적합하다.(Gartner, 2011)

2.4.1 규모(Volume)

정보통신기술의 발전과 IT의 일상화가 진행되면서 해마다 디지털 정보량이 기하급수적으로 폭증하고 있다. 빅 데이터라 불리기 위해서는 일단 데이터의 집합의 크기가 커야 하며, 그 크기가 대략적으로 적게는 수 테라바이트(TB : 10¹²=1조)에서 많게는 수 페타바이트(PT : 10¹⁵=1,000조) 정도 크기의 데이터 집합을 지칭하는 것이 일반적이다.

2.4.2 다양성(Variety)

단순히 큰 규모만으로 빅 데이터라고 할 수 없는 또 다른 이유는 오늘날 쏟아지고 있는 다양한 데이터 중 분석해야 할 데이터 형태가 매우 다양하다는 점이다. 로그 기록, 소셜, 위치, 소비, 현실 데이터 등 데이터 종류의 다양화와 텍스트 이외의 멀티미디어 등 비정형화된 데이터 유형의 다양화와 갈수록 비정형 데이터의 비율이 증가한다는 것이다.

2.4.3 속도(Velocity)

또 다른 난관은 바로 데이터가 쏟아져 들어오는 속도이다. 사물정보(센서, 모니터링), 스트리밍 정보 등 실시간 정보가 증가하고 실시간성으로 인한 데이터의 생성, 이동(유통) 속도가 증가하고 있다. 빅 데이터 시대의 기업에는 사건 발생과 데이터 수집 사이의 간격을 좁혀 더욱 역동적으로 경영 환경변화에 대응할 기회가 커지고 있다.

2.4.4 복잡성(Complexity)

구조화 되지 않은 데이터(비정형 데이터), 데이터 저장방식의 차이, 중복성 문제 등 데이터 종류의 확대, 외부 데이터 활용으로 관리대상이 증가하고, 데이터 관리 및 처리의 복잡성이 심화되어 계속적으로 이에 대응하기 위한 새로운 기법들이 등장하고 있다.

3. 빅 데이터 활용가치 및 적용사례

3.1 데이터 활용 전략

최근 글로벌 컨설팅 기관 등은 빅 데이터의 경제적 가치에 주목하고 있는데 IT의 주도권이 인프라, 기술, SW 등에서 데이터로 옮겨가고 있으며, 빅 데이터를 IT시장과 기술발전의 핵심주제로 인식하고 있다. 개인의 취향을 직접적으로 반영하는 소셜 데이터는 기업과 고객을 연결하는 쌍방향 소통 수단으로 활용되며 소셜 분석의 중요성이 증대되고 있다, 또한, 빅 데이터에서 유용한 정보를 찾아내고 활용할 수 있는 기업이 시장을 선도하는 등 새로운 경제적 가치의 원천이 될 것으로 기대하고 있다.

빅 데이터를 활용한 BI/DW 분석을 위해 현실에 기반한 기업의 실질적인 활용 전략은 당분간 기업들에게 정형 데이터와 비정형 데이터들에 대한 상호보완적인 두 트랙 분석 접근 전략이 필요할 것으로 보인다.

‘빅 데이터 시대 위한 BI/DW전략’(김희배, 한국EMC데이터 컴퓨팅 제품사업본부)에서 ‘대용량 정형 데이터를 분석하기 위해 갖추어야 하는 인프라 요구사항으로 강력한 성능, 워크로드 관리 기능, 유연성, 안정성, 확장성, 비용 효율, 개방성 등을 제시하고 있다.

3.2 데이터 활용에 필요한 3대 요소

빅 데이터 활용을 위해서는 데이터 수집부터 분석, 적용에 이르는 전 과정에 걸쳐 다양한 인프라와 조직적 역량이 필요하다. 데이터의 규모와 복잡성, 실시간

분석 속도가 기업의 경쟁력에 결정적 차이를 유발한다는 점에서 빅 데이터의 활용에서는 클라우드 컴퓨팅과 네트워크 환경, 그리고 실시간 활용성을 3대 핵심 요소로 볼 수 있다.

3.2.1 클라우드 컴퓨팅

빅 데이터는 규모가 기존의 분석 도구로 처리가 어려울 만큼 크기 때문에 클라우드 기술이 대거 활용된다. 회귀분석이나 시뮬레이션, 클러스터 분석 같은 것도 Big table이나 MapReduce, Hadoop, Hbase와 같은 클라우드 기술을 적용하여 데이터를 분석, 처리해야만 한다.

3.2.2 네트워크 인프라 구축

클라우드 컴퓨팅 기술을 이용한 실시간 분석결과의 즉각적인 집행을 위해서는 네트워크 인프라 구축도 중요해진다. 기존의 유무선 네트워크 및 주파수 인프라 관리 또한 복잡한 빅 데이터의 수집, 분석, 적용에 맞도록 재정립할 필요성이 증가한다.

3.2.3 실시간 활용성

실시간 활용성은 모바일 라이프의 확산과도 밀접하게 연계되어 있다. 사용자의 위치정보가 실시간으로 모니터링 되고 스마트 폰을 사용한 SNS가 활성화되면서 위치 정보와 사용자의 선호를 반영한 마케팅이 실시간으로 집행될 수 있는 환경이 이미 조성되어 있다.

3.3 빅 데이터 활용 영역 및 적용사례

3.3.1 빅 데이터 활용 영역

빅 데이터 분석을 통해 현재의 상황 및 가까운 미래를 예측할 수 있는 의미 있는 정보를 창출하고, 향후 비즈니스의 모습, 나아가 산업구조까지 바꿀 정도의 잠재력을 가지고 있는 것으로 평가 받고 있다. 다양한 사업, 이용자 형태분석 등에 이용할 수 있는 의미를 도출해 낼 수 있어야 하며, 빅 데이터 분석을 통해 소비자의 취향과 행동 등 다양한 변화를 실시간으로 감지하고, 사람을 통하지 않고 의사결정을 할 수 있는 경우, 그 효과는 가늠할 수 없을 정도로 높을 전망이다. McKinsey는 의료, 소매, 제조, 개인 위치정보, 공공의 5개 분야별로 빅 데이터 활용사례와 효과를 소개하였다.(McKinsey, 2011)

<표 3.1> Data utilizing overseas practices

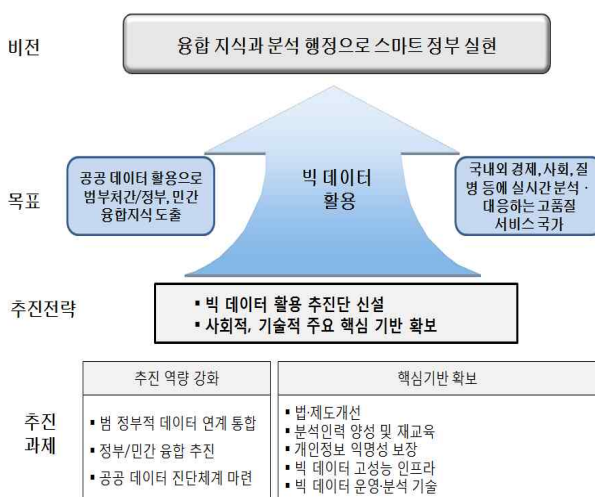
분야	활용분야	성과
의료	의료기관별 진료방법, 효능, 비용데이터를 분석하여 효과적인 진료방법 파악 가능	연간 \$3,000억 이상의 비용절감
소매업	마케팅 영역에서 교차판매, 위치기반 마케팅, 고객세분화	최소 0.5%의 생산성 향상 이윤 60%이상 증가 예상
제조업	제품개발, 공급망관리 분야 수요예측	제품개발비 20~50%절감
공공	정부 데이터베이스 공개 맞춤형 구축추천 시스템	행정비용 15~20%절감
개인 위치 정보	통신, 소매, 미디어가 결합된 영역으로 새로운 가치창출이 가능한 분야	향후 10년간 신규 서비스 공급자 매출은 \$1,000억 이상 혜택

(출처 : 정보통신정책연구원(KISDI), 정용찬, 20120,02)

3.3.2 빅 데이터 적용사례

(1) 국가 차원의 빅 데이터 활용방안 구상

국가정보전략위원회가 빅 데이터를 활용한 스마트 정부 구현 안을 제시하는 등 국가 차원의 사업 추진을 강화 중이며, ‘빅 데이터 활용 추진단이 신설되어 빅 데이터 경쟁력을 높이기 위한 국가, 사회적 기반 확보에 적극 나설 계획이다. 방송통신위원회는 2,072억원이 투자되는 '12년 R&D 시행계획을 발표하며, 빅 데이터를 주요 분야 중 하나로 지목하고 있다.(국가정보전략위원회, 2011.11, ‘빅 데이터를 활용한 스마트정부 구현)



<그림 2> Smart government plans utilize the domestic Big Data

(2) 국내 공공분야

(가) 국민권익위원회

각 부처의 온라인 민원정보를 수집, 관리, 분류, 분석 및 예측하는 민원정보 분석시스템을 2011년부터 구축 및 운영하고 있다. 월별, 지역별 주요 민원 켈런더를 제작하여 사회적 이슈를 민원 지도 형태로 제공하고, 반복 민원에 대한 대책을 마련하고 있다.

(나) 한국석유공사

고유가에 따른 소비자 부담을 줄이기 위하여 유가 정보를 예측 제공하는 오피넷 시스템을 구축하여 국제유가를 기반으로 1주일 뒤의 국내 정유사와 지역별 판매가격을 예측, 자료를 제공하여 유가에 민감한 국내 물가 안정에 기여하고 있다.

(다) 국립보건 연구원

중앙 인체자원 은행을 운영하여 한국인의 질병원인 규명, 진단, 치료법 연구개발에 필요한 인체자원 확보 및 제공을 하고 있다.

(3) 국내 민간분야

(가) 포스코

가격 변동이 큰 철광석 등 자원을 적시에 조달하기 위하여 데이터 분석을 통하여 최적 구매시기와 가격대를 결정하고 있다. 생산 공정별 온도, 습도, 압력, 성분 등의 데이터와 불량률을 결합하여 생산효율성이 높아 지도록 실시간으로 공정을 제어하고 있다.

(나) SK텔레콤

지도와 연결된 유동인구, 업종별, 월별 매출 등의 정보로 상권분석 서비스를 제공하고 있다.

(다) 다음

소셜 네트워크의 이슈와 관심 키워드의 실시간 모니터링으로 상황에 맞는 대응전략 및 마케팅 전략 수립에 지원한다.

(4) 해외 공공분야

(가) 미국

의료분야에서 약 검색을 지원하는 국립보건원의 필박스(Pillbox) 서비스로 검색 통계를 활용하여 주요 질병의 분포, 연도별 증가 등을 분석하고 연간 5천만 달러의 비용절감 효과를 발생시키고 있다.

(나) 영국

공공부문의 정보 공유 및 활용을 위한 데이터 윈스톱 서비스를 제공하고 있다.

(다) 싱가포르

테러, 전염병, 금융위기 등 국가적 차원의 위험, 기회 요인을 선제적으로 파악하고 대응하기 위한 국가위험관리시스템(RAHS : Risk Assessment and Horizon Scanning)을 2004년부터 구축하여 운영하고 있다.

3.4 빅 데이터 활용의 저해요인과 극복방안

빅 데이터 활용을 촉진하는 요인들이 증가하고 있는 반면, 저해요인 또한 존재하고 있다.

3.4.1 개인정보와 기밀정보에 대한 과제

빅 데이터 비즈니스가 진전됨에 따라 개인정보나 영업 기밀을 다루는 문제가 해결해야 할 과제로 대두되고 있다. 개인정보와 관련된 데이터로는 연령, 직업, 성별, 취미나 기호에 관한 데이터, 자산이나 건강 상태에 관한 데이터, 거주지나 연락처 혹은 콘텐츠 열람이나 구매 이력 등이 있다.

프라이버시 문제 해결을 위한 제도적 기술적 장치가 시급히 마련되어야 하며, 데이터 문제 해결을 위한 제도적 장치로는 '옵트인(opt-in)' 제도가 대표적이다. 옵트인은 이용자가 동의해야 데이터 수집 및 활용할 수 있도록 하는 제도로 안전하게 개인정보가 취급하고 있다는 사실을 이용자가 인식하도록 하는 것이다.

3.4.2 데이터 소유권 및 서비스 통제 이슈

데이터 취합 및 방법이 다양해지고, 데이터 거래 시장이 등장할 경우 데이터의 소유권 문제가 주요 이슈로 부상할 수 있다. 웹상에 축적되는 데이터의 분석, 가공을 통한 2차 데이터의 소유권 문제는 명확히 정립되어 있지 않는 상황이다. SNS를 통해 이용자가 생성한 데이터의 소유권도 문제시 될 수 있다. 블로그나 커뮤니티 서비스 등을 통해 축적된 데이터의 소유권에 대해 포털 등 서비스 업체와 이용자들 간의 분쟁사례가 벌써부터 존재하고 있다.

3.4.3 전문 인력의 부족

빅 데이터는 자료의 취합과 기계적 분석 그 자체보다는 이를 활용하기 위한 목적과 부합해 문제를 분석하고 그 결과를 해석해 실제 적용을 이끌어내는 인력의 중요성이 커지고 있다. 기존 IT에 국한된 한계적 전문인력보다는 통계학은 물론 경제학 및 인문학 등의 지식을 갖추고 문제 해결을 위한 창의적 아이디어를 도출할 수 있는 전문인력이 요구된다.

빅 데이터 산업에 필요한 전문인력은 통계와 인공지능 분야의 분석 기술과 통계, 수학, 컴퓨터 공학 등 높은 교육 및 기술수준이 필요한 분석 전문가와 분석 요청과 결과 해석을 통한 의사결정이 가능하고, 수학적 능력과 기술도 일정 수준 필요하며, 광범위한 분야의 경험을 갖춘 데이터 기반 관리자와 분석가, 빅 데이터 활용을 위한 HW와 SW 도구의 적용 및 활용이 가능한 기술적 인력지원이 필요하다.

4. 새로운 고객가치와 비즈니스 창출

4.1 새로운 비즈니스 등장 배경

기존 산업에서 제품이나 서비스가 빅 데이터와 결합하면서 새로운 가치와 비즈니스를 창출하는 기회가 디지털 시대의 자연스러운 과정으로 나타나고 있다.

빅 데이터가 몰고 올 궁극적인 미래는 지능화 또는 스마트화가 더욱 광범위하게 실현되어 사람이 구체적으로 지시하지 않아도 자동화 또는 로봇이 사람 대신 무언인가 해주는 세상이다. 빅 데이터에 의한 새로운 비즈니스 창출은 다음과 같은 기회의 확대 혹은 역량의 증가에 기인하고 있다.

첫째, 고객 소비 행동은 물론 일상적인 행동에 대해 정보를 수집함으로써 고객의 행태를 반영한 사업기회가 늘어난다.

둘째, 상황 혹은 맥락 인식 능력이 향상된다. 이는 빅 데이터에서 실시간 데이터와 상호작용 데이터의 증가와 관계가 깊다. 소비활동에 영향을 주고 받는 소비자들의 소셜 네트워크 구조를 파악하여 효과적인 마케팅을 위한 기반 정보를 확보할 수 있다.

셋째, 빅 데이터로 더욱 신속해지고 자동화된 의사결정을 활용하는 사업도 유망한 분야라고 할 수 있다. 이는 데이터 분석을 위해 다양한 종류의 엄청난 데이터를 포착하여 저장하고, 이들로부터 사람의 판단을 돕거나 대신할 정보를 자동으로 추출하는 기술의 발전 덕분에 가능해지고 있다.

넷째, 빅 데이터의 수집 및 분석을 체계화 하여 현황 파악에 걸리는 시간을 단축하고, 전략 실행 효과의 예측모형과 시나리오를 도출할 수 있다. 빅 데이터에서 관찰되는 다양한 요인이 성과에 미치는 영향을 파악하고, 이들 요인을 변화시켜가며 여러 시나리오와 기대성 과를 보다 신속하게 계산이 가능하다.

이처럼 빅 데이터를 활용한 새로운 비즈니스 중 최근 다양한 혁신으로 신 산업을 개척하고 있는 '맥락/상황인식 비즈니스'와 이를 스마트 폰 내에서 구현하는 '스마트 모빌리티 비즈니스', 그리고 '자원 이용 최적화 비즈니스' 등이 새로운 비즈니스 영역으로 등장하고 있다.

4.2 빅 데이터의 산업별 기여 분야

빅 데이터가 향후 산업 전반에 큰 영향을 미칠 것이라는 점은 모두가 예상하는 바이다. 빅 데이터에 의한 생산성 향상이 클 것으로 예상되는 산업은 정보산업, 컴퓨터/전자제품, 제조업, 금융업, 공공부문 등이며, 새

로운 가치를 창출할 가능성이 높은 산업은 정보산업, 금융업, 공공부문 등이다.(McKinsey, 산업별 빅 데이터 예상효과 분석, 삼성경제연구소, 2012.8)

4.2.1 제조업

제조업의 생산성 향상 가능성은 다른 산업들과 비교해 크지 않을 것으로 전망 된다. 제조업체들은 이미 ERP나 SCM 등 기존 정보기술로 상당 부분 생산성이 향상되고 있다.

4.2.2 금융업

금융업은 그 자체가 정보산업으로서 다른 어느 산업보다 빅 데이터의 영향이 클 것으로 전망한다.

각종 위험관리와 마케팅에서 빅 데이터 경영혁신이 기대되는 산업도 금융업이다. 앞으로 유통업 등 다른 산업과 금융업의 융합이 가장 활발한 분야로서 빅 데이터에 의한 새로운 사업의 탄생을 빈번해질 것으로 전망된다.

4.2.3 ICT산업(IT+통신 산업)

ICT산업은 빅 데이터의 실험장이 될 가능성이 높은 산업이다. 모든 빅 데이터 경영혁신 분야에서 ICT 산업은 다른 산업들을 선도할 것으로 예상된다. 특히 통신업체들과 인터넷 기업들은 인프라 산업에서 벗어나기 위해 빅 데이터 분석으로 다른 업체들과 차별화된 서비스를 창출하고 새로운 수익원을 발굴하는 것이 앞으로의 핵심이 될 전망이다.

4.2.4 소매 유통업

소매 유통업은 유행에 민감한 사업으로 그 어느 분야보다 고객의 수요와 취향 변화를 빨리 파악하고 대응하는데 빅 데이터를 중요하게 활용한다. 의사결정의 스마트화가 결국 유통업의 경쟁력을 결정할 것으로 예상된다. 또한 유통업은 수많은 외부 협력업체들과 공급사슬관리로 연결되어 있어 유통업의 생산성 향상은 선택이 아닌 필수 과업이다.

4.2.5 공공부문

현재 정부 각 부처에서 국민들에게 빅 데이터를 활용한 각종 민원 및 행정 서비스를 경쟁적으로 계획하고 있듯이 공공 부문에서 새로운 고객가치를 창출하는데 빅 데이터를 활용하는 경향은 전 세계적인 흐름이 될 전망이다.

5. 빅 데이터 확산에 따른 도전과 기회

소셜 미디어에서 촉발된 빅 데이터 이슈가 산업 전반으로 확산되면서 새로운 가치와 수익 창출원으로 기대를 모으고 있으나, 빅 브라더에 대한 우려와 전문인력 부족 등 해결해야 할 많은 도전과제도 생존하고 있다.

빅 데이터에 대한 일반적인 정의는 아직 합의되지 않았지만 맥킨지 글로벌 연구소(MGI)에 따르면 빅 데이터를 ‘전형적인 데이터베이스 도구의 과잉화, 저장, 관리 분석 역량을 넘어서는 크기의 데이터 집합’이라 정의하고 있다.

양적 측면의 거대함을 의미하던 빅 데이터는 최근에 점차 관련 도구, 플랫폼, 분석 방법론까지 포함하는 용어로 변모하고 있다. 빅 데이터는 정형화된 데이터와 달리 텍스트, 사진, 동영상 등 포맷이 제각각으로 비정형적인 특성을 가지고 있으며, 따라서 빅 데이터가 기존 분석도구의 역량을 넘어서는다는 말은 정형화 데이터를 상징하고 있는 기존 분석 경영의 기술로는 활용하기 적합하지 않다는 의미도 포함하고 있다. 이런 이유로 최근 빅 데이터는 기존의 구조화 데이터 분석도구와는 다른 방식, 다른 분석 방법으로 대용량 데이터를 저장, 수집, 발굴, 분석, 비즈니스화 하는 일련의 과정을 포괄하는 말로 바뀌고 있다.

빅 데이터의 핵심은 바로 데이터 비즈니스화에 있으며, 막대한 양의 1차 데이터가 아닌 가공된 2차 데이터가 진정한 의미의 빅 데이터이다. 빅 데이터는 물이나 공기처럼 흘러넘치는 데이터를 빅 머니(Big money)로 변환시키는 것이며, 하드웨어에 갇힌 데이터가 아닌 소프트웨어처럼 부가가치를 창출하는 도구로서 데이터를 활용하는 의미이며, 데이터가 필요한 사람에게 엄청난 양의 데이터를 쏟아 붓는 것이 아니라 적당하고 적절한 양의 데이터를 제공하는 것이 바로 빅 데이터의 역할이며, 이런 역할이 새로운 비즈니스를 만들고 새로운 수익을 창출하게 된다.

6. 시사점

본 연구에서는 기업이 적극적으로 빅 데이터를 도입해야 하는 다양한 이유와 활용 방향, 국내외 선진 적용 사례 등을 살펴보았다.

데이터는 산업혁명의 석탄과 같이 스마트 시대의 중요한 ‘자원’임을 인식하고, 이를 효과적으로 관찰할 수 있는 정책적, 조직적 대책마련이 필요하다. 특히 인터넷(네트워크)은 스마트 생태계에서 ‘자원(데이터)’이 생산되는 원천이므로 새로운 데이터 생태계의 발전은

미디어 정책과 매우 밀접한 관련이 있다. 따라서 빅 데이터 혁명에 효과적으로 대응하기 위해서는 데이터의 생산, 유통, 소비 전 영역에서 패러다임의 전환이 필요하다. 이제 빅 데이터는 대중화 초기 단계로 진입하여 관련기술이 발달하고 민간기업의 수요가 꾸준히 증가하여 곧 대중화 단계로 접어들 것으로 예상된다. 이러한 주요 요소들을 반영하여 빅 데이터 활용 원칙을 몇 가지로 제시해 본다.

첫째, 빅 데이터의 활용은 일종의 기업 문화다.

기본적으로 조직 내에서 빅 데이터를 활용하기 위해서는 그 필요성과 가치에 대해 공감대가 형성되어야 한다. 즉 조직의 전략적 자산으로 빅 데이터를 활용하고, 경영진이 적극적인 자원을 아끼지 않으며, 빅 데이터로부터 나오는 지식에 전사적으로 접근할 수 있어야 빅 데이터가 기업의 주축문화가 된다.

둘째, 정보관리를 실현하고 적절한 사람을 확보 및 배치하여 정보관리 역량을 확보한다.

모든 조직원이 데이터를 신뢰할 수 있어야 한다. 믿을 수 있는 유용한 데이터가 존재하고, 필요한 데이터가 어디에나 있으며, 그 데이터를 타 부서가 쉽게 공유할 수 있는 시스템은 기본적인 정보관리의 출발점이다. 셋째, 적절한 과제를 선택해 실행경험을 축적하는 것이다. 데이터의 활용을 가장 큰 기회에 먼저 집중하는 것이 필요하다. 생산된 지식을 활용하고, 활용 사례를 홍보해야 한다.

넷째, 기술도 중요하지만 시작이 더욱 중요하다.

기존에 보유한 접근 가능한 데이터들의 분석마저도 본격적으로 시작하지 않은 기업들이 많은 현실에서 최신 기법과 기술을 도입하는 것만이 빅 데이터라고 생각하는 것은 오산이다. 기존 데이터를 분석적으로 활용하는 것도 빅 데이터의 시작이다.

분명 빅 데이터의 역량을 얻고 원활히 활용하기까지의 여정은 길고 고단할 것이다. 하위 유형에서 출발하여 거의 '전략적 데이터 관리자' 유형까지 올라와 성과를 내고 있는 기업들의 사례를 분석한 결과를 살펴보면 데이터 분석을 시스템화하고 조직에 완전히 정착시켜 성과를 끌어 내기까지 무려 평균 8년이라는 걸렸다고 한다. 앞으로 많은 기술과 산업 환경이 어떻게 바뀌고 전개될지 모르는 기술의 변화 및 시장과 사회의 수용 흐름까지 감안한다면 10년 이상을 내다보는 안목과 인내가 절실히 필요하다 하겠다.

7. 참고 문헌

- [1] SERI(2012), "Bid Data, Change Management", Hamyugeun, Chaeseungbeom
- [2] Suzuki, Ryosuke, "Big Data Business", The Forest
- [3] SERI(2011), "Find gold mines in the Data Deluge : 'Big data analysis and use', Chaeseungbeom
- [4] Leesungchoon(2012), "Big Data utilizing and Implications for Telecommunications Industry"
- [5] Gimhuiuae(2011), "Big Data era for BI / DW strategy", Network Times
- [6] National ICT Strategies(2012), "Big Data Master Plan Current Status and Future Plans", Ministry of Public Administration and Security
- [7] National Information Society Agency(2011), "New value creation engine, new possibilities in Big Data and Strategies", IT& Future Strategy
- [8] National IT Industry Promotion Agency(2012), "Latest IT trends - challenges and opportunities due to the proliferation of Big Data"
- [9] Korea Internet Agency Policy Planning Team (2012), "Big Data era, new possibilities and challenges", Attleruss Research & Consulting
- [10] Jeongyongchan(2012), "Big Data revolution and media policy issues", Information and Communication Policy Research Institute

저자 소개

고준철



현재 AB&S(주)컨설팅 사업부 이사로 재직 중이고 IT산업분야의 컨설팅 서비스와 교육훈련 강사로 활동하고 있으며, 주요 관심분야는 IT Governance, ISP/BPR, 정보보안경영시스템, 소프트웨어 프로세스 개선(SPI), 소프트웨어공학 분야로 명지대학교에서 공학박사학위를 취득하였다.

주소: 경기도 안양시 만안구 안양6동 호정타워 913호

이해육



현재 육군본부에 재직하고 있으며, 명지대학교에서 산업시스템공학 석사 및 박사과정 중. 관심분야로 군사시설이 인근 주민과 지자체에 미치는 양면영향과 관계개선 연구, 민·관·군 상생방안 등을 연구 하고 있다.

주소 : 경기 고양시 덕양구 화정2동 별빛아파트 918/606

정지윤



한중법무법인 국제법무팀 이사와 명지대학교 산업대학원 국제교류경영학전공 겸임조교수로 재직 중이며, 법무부 법무정책 온라인 국민평가단과 (사)한국다문화센터 운영위원으로 서비스품질, 컨설팅, 국제교류경영, 이민/유학/해외취업/다문화/출입국관련 교육 등에 활동 중이다. 영어영문학학사와 이민학석사와 산업시스템경영학석사와 현재 명지대학교 산업경영공학과 박사과정에 재학 중이다.

주소: 인천 중구 운서동 2709-3 주공스카이빌 APT.705동 301호

강경식



인하대학교 산업공학과에서 학사·석사·박사와 연세대학교·경희대학교에서 경영학 석사·박사 취득. North Dakota State Univ.에서 Post -Doc과 Adjunct Professor 역임. 현재 명지대학교 산업경영공학과 교수로 재직 중. 주요 관심분야는 생산관리, 물류관리, 안전경영 등이다.

주소: 경기도 용인시 처인구 남동 산 38-1 명지대학교 산업경영공학과