

〈튜토리얼〉

빅데이터, 비즈니스 애널리틱스, IoT: 경영의 새로운 도전과 기회

장 영 재*

〈목 차〉

I. 빅데이터와 비즈니스 애널리틱스	2.2 사물인터넷의 과학적 의미
1.1 빅데이터의 개념적 정의	2.3 사물인터넷 활용 사례
1.2 빅데이터 개념 탄생의 배경	2.4 빅데이터-사물인터넷 혁신사례
1.3 비즈니스 애널리틱스	III. 정리
1.4 데이터 기반의 의사결정 플랫폼	참고문헌
II. 사물인터넷	<Abstract>
2.1 사물인터넷의 발전	

I. 빅데이터와 비즈니스 애널리틱스

1.1 빅데이터의 개념적 정의

90년대 인터넷 기술이 대중의 삶에 들어오면서 정보통신의 새로운 혁신이 시작되었다. 이후 컴퓨팅 기술의 발달과 저장매체 가격의 급속한 하락으로 다양한 데이터를 처리하고 보관할 수 있는 세상이 열리게 되었다. 이러한 방대한 데이터가 쌓이면서 데이터를 분석하여 새로운 인사이트를 찾거나 데이터를 통해 과거 불가능했던 새로운 혁신을 가능하게 한 개념을 빅데이터라 칭한다. 아직 학계에서는 이러한 빅데이터의 정의에 대한 다양한 의견들이 존재한다 [1]. ‘빅데이터’란 이름의 조직을 신설한 기업 및 정부 기관들조차 빅데이터의 실질적 업무를 명확하게 규정하지 못하는 것은 빅데이터란 단어의 태생이 기술이 아닌 가치에서 탄생했기 때문이다.

실제 빅데이터란 단어는 2000년대 야후 와 구글과 같은 인터넷 검색엔진 서비스를 제공하는 기업과 대용량데이터 인프라를 제공하는 IT기업에서 기존 데이터 양과 비교될 수 없는 양의 데이터를

* KAIST 산업및시스템공학과 조교수, yjang@kaist.ac.kr

보관, 처리, 분석하는 방식을 고민하며 생긴 개념적 용어다 [2]. 여기서 기술과 개념의 차이를 좀더 명확하게 예를 들어 구별하자면 최근 기계학습법 혹은 머신러닝 (machine learning)에서 주목받는 딥러닝 (deep learning)은 분명히 기술을 의미한다. 이론적 배경이 존재하고 이론을 기반으로 구체적인 활용 방법론도 존재한다. 하지만 빅데이터는 특정한 기술이나 이론이 아닌 쌓여가고 늘어가는 데이터를 좀더 고차원적으로 분석하기 위해 필요한 다양한 기술의 범위와, 데이터를 활용해 창출하는 새로운 가치를 의미한다. 빅데이터에서 기술적 요소를 찾자면 대용량 데이터를 처리하는 데이터 베이스나 데이터 관리, 데이터를 분석해서 인사이트를 찾는 통계나 데이터 마이닝 그리고 기타 데이터를 통해 좀더 나은 의사결정이나 관련 알고리즘을 기술적 범주에 포함할 수 있다. 이처럼 다양한 기술을 포괄적으로 수용하는 개념으로 빅데이터를 해석할 수 있다.

“빅데이터 전문가”라면 데이터관련 다양한 전문가들을 포괄적으로 칭하는 것이다. 이는 과거 20년 전 인터넷 기술이 일반인들의 삶에 들어올 무렵 “인터넷”을 정의하는데 학자들 사이 많은 논쟁이 있었던 것과도 유사할 수 있다. 인터넷이란 기술 자체는 이미 70년대 통신기술로 존재했었지만 인터넷이 정보통신의 혁신의 단초를 마련한 것은 인터넷 브라우저인 넷스케이프 (Netscape)가 등장하면서 부터다. 그럼 인터넷 전문가란 인터넷이란 특정 통신 기술자를 의미하는지 인터넷 공간에서 정보 공유와 활용의 편의성을 제공하는 브라우저 개발자인지 혹은 인터넷 공간상의 데이터를 처리하는 기술자를 의미하는지 모호하게 된다. 즉 인터넷이란 단어의 정의는 특정 통신기술을 지칭하는 것이었지만 인터넷을 통한 정보 혁신에는 이러한 통신 기술뿐만 아닌 다양한 기술 및 가치가 내제되어 있다 할 수 있다. 빅데이터 또한 다양한 기술과 가치의 집합적인 개념이라 정의하는 이유도 이와 유사하다.

1.2 빅데이터 개념 탄생의 배경

앞장에서는 빅데이터를 특정 기술이 아닌 개념이라 정의하였다. 이러한 개념의 탄생 배경을 이해한다면 그 개념의 의미를 이해하는데 도움 될 것이다. 앞장에 빅데이터란 용어는 2000년대 야후와 구글과 같은 인터넷 검색엔진 서비스를 제공하는 기업과 대용량데이터 인프라를 제공하는 IT기업에서 기존 데이터 양과 비교될 수 없는 양의 데이터를 보관, 처리, 분석하는 방식을 고민하며 생긴 개념이라 소개하였다.

그렇다면 인터넷 검색엔진 서비스에서 다루는 데이터의 특징이 무엇이길래 이러한 새로운 개념이 나오게 된 것일까? 검색엔진 데이터의 특징은 데이터의 실시간성 (real-time), 규모 (volume), 그리고 이질성 (variety)로 정의 할 수 있다. 실시간으로 유입되는 질의 (query)를 실시간으로 분석해 그 결과를 제공해야 함으로, 실시간으로 데이터를 처리하고 분석하는 기능은 검색엔진 서비스의 기본 요구사항이라 할 수 있다. 다음은 규모이다. 여기서 규모란 절대적인 데이터양 (stock-volume)을 의미하지 않고 유입량 (flow-volume)이라 생각하면 이해하기 쉽다. 실시간으로 들어오는 엄청난 량

의 데이터를 처리하고 분석하는 기술이 필요하다. 그리고 이러한 데이터는 지속적으로 쌓이게 될 것임으로 데이터가 증가하더라도 분석이 가능한 알고리즘이 요구되고 있다.

마지막으로 다양성은 데이터의 형태라 할 수 있다. 검색 서비스는 실시간으로 생성되는 다양한 형태 - 동영상, 이미지, 음성, 텍스트 - 의 데이터를 인식하고 저장하는 기술이 필요하다. 즉 이러한 실시간성, 규모, 이질성을 다루는 데이터 저장기술, 분석기술, 실시간 처리 능력 등이 검색엔진 서비스에서 활발히 연구되고 이러한 특징을 다루는 기술의 통합적 (holistic)인 개념이 빅데이터란 단어로 함축되어 통용되기 시작하였다.

1.3 비즈니스 애널리틱스

그렇다면 왜 이러한 검색엔진 분석에 사용된 개념이 사회/문화적으로 이슈가 되기 시작한 것일까? 2011년 이후 빅데이터란 개념은 기술 분야뿐만 아닌 사회, 문화, 정치 등 우리 삶의 이슈이자 혁신적 패러다임으로 부각되고 있다. 그렇다면 왜 이 기술 중심적 단어가 혁신적 패러다임의 이슈로 증폭되었을까? 마치 인터넷이란 개념이 수십 년간 소수의 네트워크 전문가들만의 기술에서 90년대 세상을 바꾸는 혁신적 기술로 부상한데는 인터넷 브라우저 기술이 등장했기 때문이다. 그럼 빅데이터 개념이 확산된 배경은 무엇일까?

여기에 바로 “비즈니스 애널리틱스 (Business Analytics)”가 있다. 이 단어는 2003년 미국 뱁슨 대학 (Babson College)의 토마스 데이븐포트 (Thomas Davenport)교수가 2007년 출간한 “애널리틱스로 경쟁하라 (Competing on Analytics)¹⁾”란 책의 제목에서 유래한다. 저자는 책에서는 2000년대 후반 각 산업분야의 리더 기업들의 특징을 분석하였으며 그 결과 이들 기업들은 방대한 데이터를 적극 활용하여 혁신한 공통점이 있다는 사실을 발견하였다. 기존 기업에서 활용한 데이터 분석방식이 아닌 검색엔진 기술자들이 고민하는 데이터의 실시간, 규모, 이질성에 대한 이슈를 이들 기업들도 함께 고민하고 있었으며 검색엔진 기술자들이 사용하는 다양한 기술과 알고리즘이 기업현장에서 상용되고 있음을 소개하였다. 이 책에서 이러한 고차원 기술을 다루는 기술자들을 데이터 과학자 (data scientist)라 정의했고 이러한 데이터 과학자는 단지 검색엔진 서비스를 제공하는 기업에만 있는 것이 아닌 다양한 기업에서도 활약할 수 있고 이미 각 산업별 선두기업에서는 적극적으로 이들을 활용하고 있다는 사실을 소개하였다. 전통적인 유통 분야의 월마트, 인터넷 상거래 분야의 아마존닷컴, 엔터테인먼트 분야의 넷플릭스 등을 소개하며 이들 기업들은 각 분야에서 데이터를 기반한 혁신적 운영과 서비스를 제공하고 있으며 컴퓨터 공학, 통계학, 산업공학, 수학 및 기타 공학을 전공한 전문가들이 데이터 과학자로 혁신을 주도한다는 사실을 강조하였다 [3].

1) Davenport, Thomas H., and Jeanne G. Harris. Competing on analytics: The new science of winning. Harvard Business Press, 2007. (국내에서는 <분석으로 경쟁하라>라 소개되었지만 여기서 분석이란 단어가 일반적인 분석과는 다소 차이가 있어 “애널리틱스로 경쟁하라”라고 본 논문에서는 칭하였음)

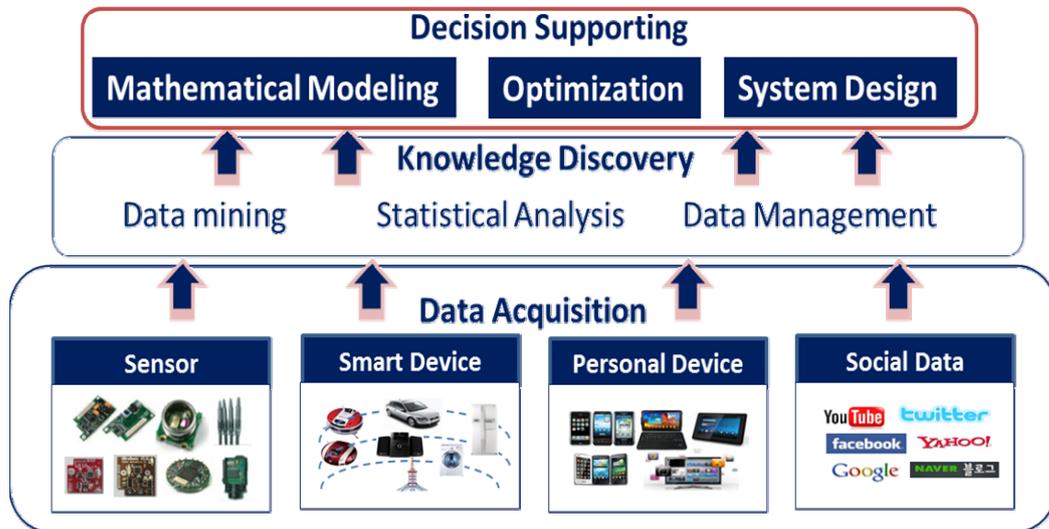
토마스 데이븐포트는 이 책에서 이러한 애널리틱스는 기존 데이터 분석 (data analysis)과의 차별점을 분명히 하고 있다. 즉 기존 분석은 정형화된 데이터, 그리고 과거 제한적인 데이터를 기반으로 현상을 파악하는 탐험적 분석 (exploratory analysis)라면 데이터 과학자들은 실시간으로 유입되는 대용량 데이터를 자신의 목적에 맞게 처리하여 향후 상황을 분석하는 예측형 애널리틱스 (predictive analytics)를 지향하고 있으며 실시간 의사결정을 지원하는 알고리즘 개발에도 적극적이라 역설하였다 [4]. 즉 빅데이터의 개념을 비즈니스의 운용에 활용하여 비즈니스를 혁신한다는 개념이 바로 비즈니스 애널리틱스란 것이다.

2011년 초 MIT 슬론 매니지먼트 리뷰 (Sloan Management Review)에는 빅데이터와 빅데이터를 분석하는 기술인 애널리틱스가 기업의 경쟁력이 될 것이라던 커버스토리를 통해 데이터가 단순한 트렌드 분석이나 보고서 작성 등의 부차적인 지원이 아닌 기업 역량의 핵심이 될 것이라 소개하며 경제와 비즈니스의 빅데이터의 활용을 구체적으로 다루었다 [5].

1.4 데이터 기반의 의사결정 플랫폼

빅데이터 개념이 어떻게 비즈니스 현장에 활용되며 혁신을 주도할 수 있는지 알아보자. <그림 1>은 데이터 기반의 의사결정 플랫폼을 설명하고 있다. 그림에서 가장 하부단은 데이터를 수집하는 하드웨어인 센서기술 및 저장기술 등으로 이뤄진 “데이터 수집 (Data Acquisition)” 단계다. 여기서 센서란 물리적인 센서뿐만 아닌 다양한 데이터 수집하는 장치를 의미한다. 이러한 센서들이 취합한 데이터를 저장하고 관리하는 기술 및 이들의 데이터에서 취합된 정보를 기반으로 데이터 마이닝 기법을 통해 새로운 인사이트 (Insight)를 얻는 것이 중간단계인 “지식의 발견 (Knowledge Discovery)”이다. 새롭게 발견된 지식을 기반으로 최적의 의사결정을 내리는 것이 가장 상위의 “의사결정 (Decision Making)”단계라 할 수 있다. 기존의 의사결정 플랫폼인 비즈니스 인텔리전스 (Business Intelligence - BI)는 복잡하고 다양한 정보를 의사결정자가 인식하기 용이하게 시각화를 통해 전달하는 것이라면, 최근의 의사결정 플랫폼은 복잡하고 다양한 데이터를 기반으로 최적의 의사결정을 수학적 알고리즘에 의해 자동으로 판단해 주는 것이라 할 수 있다.

맨 아래 단의 데이터 수집은 기존 ERP나 IT시스템을 통한 데이터 수집 및 관리 뿐만 아닌 소셜미디어, 각종 센서, 영상 정보 등 다양한 데이터를 수집하고 관리하는 기능이다. 이러한 기술은 과거 기업에서 IT시스템과 같은 정보 인프라를 다루는 기술에서 한 차원 더 나아가 늘 바뀌는 데이터 환경에서 유연하게 대처하는 기술을 요구하고 있다. 즉 과거에는 ERP나 기업정보 시스템을 셋업하고 관리하는 역할이었다면 이제는 새로운 형태의 데이터 수집방식이 필요하다. 그리고 이러한 데이터를 수집 관리할 수 있는 리소스를 제공하고 이를 기존 데이터와 연결하여 신속히 분석을 진행할 수 있도록 대처하는 능력이 필요하다. 최근 빅데이터가 이슈가 되면서 이중데이터를 다루는 데이터베이스가 등장하는 이유도 데이터 수집단의 역동성이 증가했기 때문이다.



<그림 1> 데이터기반 의사결정 플랫폼 (Data Driven Decision Platform - D3P)

<그림 1>에서 지식의 발견 (knowledge discovery) 부분은 아래 단에서 취합된 데이터를 기반으로 새로운 의미를 찾는 것이 목적이다. 통계적인 분석, 데이터 마이닝, 기타 데이터 분석 기능이 이 부분에 속한다. 하지만 과거 통계나 데이터 마이닝의 기술적 영역은 ‘누군가 데이터를 쌓아주고 잘 정리해 준다면’ 란 가정아래 분석만 담당했다면 최근 빅데이터 이슈에서는 분석가가 원하는 행태로 데이터를 모으고 취합하며 관리하는 영역으로도 확장되고 있다. 또한 반대로 데이터를 모으고 관리하는 전문가들이 분석의 영역으로도 기능을 확대하고 있다. 이는 과거 데이터 관리는 IT 부서에서, 데이터 분석은 현업 부서에서 진행하던 틀에서 벗어나 영역이 통합되고 있음을 의미한다. 실제 IBM에서 지난 3년간 빅데이터 솔루션을 CIO와 CMO 즉 최고 IT 책임자와 최고 마케팅 책임자가 함께 일하는 것이 빅데이터라 이야기한 것과 일맥상통하다 [6]. 국내에서도 신한카드의 빅데이터센터와 코오롱 스포츠의 빅데이터 분석팀이 과거 IT기술팀과 마케팅 분석팀을 통합해 새로운 빅데이터 부서를 만든 것도 이러한 이유다.

하지만 아무리 좋은 분석을 통해 새롭고 가치있는 인사이트를 제공해도 분석의 결과가 현상을 분석하거나 리포트로만 제공된다면 그 역할은 언제나 제한적일 수 밖에 없다. 분석의 목적은 분석을 통해 현업에서 이를 적용하고 새로운 혁신이나 명확한 결과를 제공하는 것이다. 리포트 자체가 목적이 아닌 행동하여 결과를 도출하는 것이 목적이 되어야 한다. 예를 들어 어느 기업에서 “매장을 찾는 고객의 80%가 연소득 3천만 원의 30대 미혼여성이다”란 인사이트를 찾았다면 이를 기반으로 새로운 서비스나 방안을 마련하는 등 구체적인 행동이 필요하다. 이러한 행동을 취하기 위해서는 매우 구체적인 의사결정이 이뤄져야한다. 즉 이들 고객들을 지속적으로 유치하기 위해 쿠폰이나

할인 제공한다면 얼마나 자주 어떤 방식으로 어떤 제품을 몇 퍼센트를 할인해야 하는지 등등의 구체적인 의사결정이 함께 이뤄져야 한다. 더구나 분석의 책임 주체가 아닌 현업부서에서는 빅데이터 부서나 기타 분석전담 부서에서 찾은 인사이트만 가지고 특정 행동을 취하는 것에 두려움이 있다는 사실은 이미 많은 연구결과에서 증명하고 있다 [7]. 결국 데이터를 분석한 빅데이터나 분석 전담 부서에서 현업과 함께 구체적인 행동안을 도출하고 의사결정을 구체화할 필요가 있다. <그림 1>에서 가장 맨 상위 단계 데이터에 기반한 실시간 의사결정 알고리즘 부분이 존재하는 이유다.

데이터 기반 의사결정이라 함은 분석에서 도출한 인사이트를 기반으로 행동에 옮길 때 필요한 구체적인 의사결정을 고도화된 알고리즘을 통해 구체화 한 것이다. 예를 들어 MIT공대와 패션브랜드 Zara가 2000년대 초반 진행한 프로젝트에서는 실시간으로 유입되는 전 세계 매장의 판매 정보 및 기업 내 다양한 정보를 기반으로 단기 수요예측을 진행하고 이를 기반으로 어느 매장에 어떤 제품의 어떤 사이즈가 얼마만큼 공급되어야 하는지를 정수혼합계획법 (Mixed Integer Linear Programming)으로 결정한다 [8]. 카이스트와 코오롱 스포츠도 2014년 전국 200여 개 매장의 판매 트렌드를 분석하고 이를 바탕으로 매장의 구체적인 실시간 배분 계획을 최적화 알고리즘으로 구현해 현장 운영을 혁신한 사례가 있다 [9]. 그 외 아마존닷컴의 실시간 연관 제품 추천, 넷플릭스의 영화 추천 서비스 등도 데이터에 기반한 의사결정 알고리즘을 제공하는 사례라 할 수 있다 [10].

데이터 기반 의사결정 플랫폼에서 데이터의 가치 흐름은 단순히 아래서 위로만, 즉 센서로부터의 데이터 수집-분석-의사결정 순으로만 이뤄지지 않는다. 오히려 데이터를 기반으로 새로운 가치를 창출할 때는 어떤 의사결정이 필요하고 이러한 의사결정을 위해서는 어떤 인사이트가 필요하며 이를 위해서는 어떤 데이터가 필요한지 상위에서 하위 역순으로 필요한 프로세스, 정보 시스템, 그리고 데이터가 정의되기도 한다. 즉 데이터가 먼저인지 의사결정이 먼저인지는 의미가 없으며 수집된 데이터를 바탕으로 새로운 의사결정을 내리기도 하고 의사결정을 내리기 위해 새로운 데이터를 수집하기도 한다 [10].

II. 사물인터넷

2014년 인터넷에서 유통되는 정보의 상당량이 사람의 의해 창조된 정보가 아닌 사물이 정기적으로 자동으로 수집한 데이터이다²⁾. 여기서 사물 (things)이란 자동 정보 기기 및 센서를 의미한다. 이러한 대용량 데이터의 원천생성 역할을 하는 센서와 인터넷과의 네트워크를 통해 전달되는 데이

2) SAS Institute, Gartner, Deloitte Analytics 등 다양한 분석 및 데이터 기반 기업들이 인터넷 정보량 (traffic)에서 사물 (센서)에서 생성된 량을 추정할 보고가 있지만 기관마다 산정한 가정과 방식이 상이하고 구체적인 수치를 측정하는 것은 불가능하고 아직 학문적으로 그 량을 과학적으로 입증한 연구는 소개된 적이 없다.

터는 빅데이터에서 언급하는 데이터의 특징을 지니고 있다. 이러한 의미에서 사물인터넷 (IoT)은 빅데이터를 논하는데 빠질 수 없는 주제이다. 본 장에서는 사물인터넷의 의미와 빅데이터 관점에서의 그 가능성을 파악하도록 한다.

2.1 사물인터넷의 발전

CCTV에서 생성된 영상정보, 택시에서 생성되는 결제정보, 기상청에서 실시간으로 날씨정보를 제공하기 위해 설치된 온도와 습도 센서와 같은 정보도 네트워크를 통해 인터넷으로 전달된다. 이처럼 사물인터넷은 이미 인터넷이 처음 개발될 당시부터 제안되고 활용된 방식이다. 단지 당시에는 센서가 컴퓨터와 연결되어야만 가능했다면 최근에는 센서 자체가 통신 및 컴퓨팅역할을 수행하는 독립적인 개체로서 지능화되었다는 것이 차이일 뿐이다.

이미 70년대 인터넷 개념이 확립될 당시에도 사물인터넷은 존재했지만 최근 들어 더욱 사회의 주목을 받는 이유로는 스마트폰의 대중화와 무관하지 않다. 자체 무선네트워크 통신기능을 가진 스마트폰이 몸에 착용하는 웨어러블 기기 및 센서와 통신하면서 손쉽게 센서 정보를 인터넷으로 전송할 수 있게 되었다. 몸의 생리 및 활동 데이터를 취합해서 인터넷으로 공유하는 기기에서부터 차량 OBD (On-Board Diagnostics) 단자에 접속하여 다양한 데이터를 휴대폰을 통해 인터넷으로 공유하는 장비까지, 생활의 다양한 정보를 실시간으로 취합하고 이를 인터넷으로 전달하고 공유하는 시대에 접어들었다.

하지만 사물인터넷의 대중화에 가장 큰 기여를 한 것은 무엇보다도 스마트폰 대중화로 인한 다양한 센서의 가격 하락이라 할 수 있다. 스마트폰에는 고화질 소형 카메라, GPS 센서, 가속센서 (Accelerometer), 자이로 센서 (Gyro sensor) 등 첨단 센서들이 탑재되어있다. 스마트폰 대중화가 이뤄지기 전 이러한 센서들은 첨단 계측장비나 특수 장비 목적으로만 활용되어 상당한 고가에 판매되었다. 하지만 스마트폰 대중화는 이러한 센서의 대량생산을 이끌어냈고 이로 인해 급속한 가격하락이 이뤄졌다. 예를 들어 고가 계측 장비로만 판매되던 가속센서 및 자이로센서는 10년 전만해도 판매가가 개당 수만 달러에 달했지만 최근에는 수천 달러 미만으로도 최고급 센서를 구매할 수 있고 일반 스마트폰에 탑재된 자이로 센서와 가속센서는 가격에 몇 달러 미만으로 스마트폰 제조사에 공급되고 있다³⁾.

2.2 사물인터넷의 과학적 의미

그렇다면 이러한 센서기술의 발전 및 하락 그리고 이러한 센서의 스마트 기기와의 연동이 어떤

3) 테크 인사이트에서는 삼성 갤럭시 S6내의 자이로센서와 가속센서의 가격을 \$4.50대로 추정하고 있다 (<http://www.techinsights.com/teardown.com/SamsungGalaxyS6/>)

의미를 지니는지 알아보자. 이러한 센서 기술 발전 및 가격하락은 과거 측정하지 못했던 새로운 계측을 가능하게 한다. 예를 들어 과거 스키 선수들의 지도는 지도자의 시각적 관찰에만 의지하였다. 즉, 지도자는 학습자인 선수의 몸동작을 직접 관찰이나 비디오 관독으로만 지도할 수 있었다. 물론 몸의 움직임을 관찰할 수 있는 광학장비는 이미 보편화되어 실내 스포츠에서는 이러한 장비를 통해 몸의 움직임을 데이터화 한 후 선수 지도에 활용하는 방법은 스포츠 선진국에서 보편화 된 방식이다. 하지만 스키와 같이 외부에서 큰 외부 공간을 활용해 빠른 속도로 몸의 움직임이 이뤄지는 운동에서는 광학장비를 통한 동작분석은 불가능했다. 그러나 최근 해외 몇몇 업체에서 과거 자동차나 차량에 탑재된 고성능 자이로 센서와 가속센서를 인체에 부착할 수 있는 크기로 소형화에 성공하였고 대량 생산을 통한 대중화를 이뤄 학교 실험실에서도 연구 목적으로 활용할 수 있는 길이 열렸다. 과거 불가능했던 스키 선수의 활강 데이터가 실시간으로 취합되고 이를 통해 새로운 운동역학 분석의 길이 열린 것이다 [11].

자이로 센서와 가속센서와 같은 첨단 센서로 선수의 경기력 향상을 도모한 사례는 2014년 월드컵 우승을 차지한 독일 국가대표팀의 사례에서도 찾을 수 있다. 독일 대표팀은 독일의 대표적인 IT기업인 SAP의 후원으로 선수 개인의 동작과 움직임 그리고 팀의 전체 움직임을 센서로 파악하여 개별 맞춤형 훈련 프로그램을 개발하였고 전력 분석 및 구상에도 데이터를 활용하였다. 국내에는 스포츠 빅데이터 승리로 알려졌지만 이러한 성공 스토리 이면에서는 센서-데이터분석-의사결정의 <그림 1>에서와 같은 의사결정 플랫폼 기반으로 결과 도출된 것이다.

2.3 사물인터넷 활용 사례

그렇다면 산업에서의 사물 인터넷 활용의 의미는 무엇일까? 일단 직접적으로 과거에는 불가능했던 새로운 정보를 취합할 수 있다는 점에서 그 가치를 해석할 수 있다. 앞서 스키의 사례와 같이 사람의 동적인 움직임을 기반으로 작업자의 상황을 파악할 수 있다. 예를 들어 선박 제조나 대형 구조물 제조에서 각 작업자들의 현재 위치 파악과 몸에 부착된 다양한 센서를 통해 작업자의 현재 진행 작업을 파악할 수 있다. 이와 비슷한 사례로 볼보 (Volvo) 상용차는 포크레인, 불도저와 같은 건설 중장비에 다양한 센서를 달아 장비의 작업을 실시간으로 파악해서 작업 진척도와 작업 계획을 수립하는 모델을 개발하였다. 토목이나 건설 공사를 공장으로 비유하자면 건설 중장비는 공장의 기계로 비유할 수 있다. 즉 장비의 다양한 센서 정보로 어떤 작업이 언제 이뤄지고 이를 기반으로 작업 계획 수립 및 관리를 한다는 개념이다.

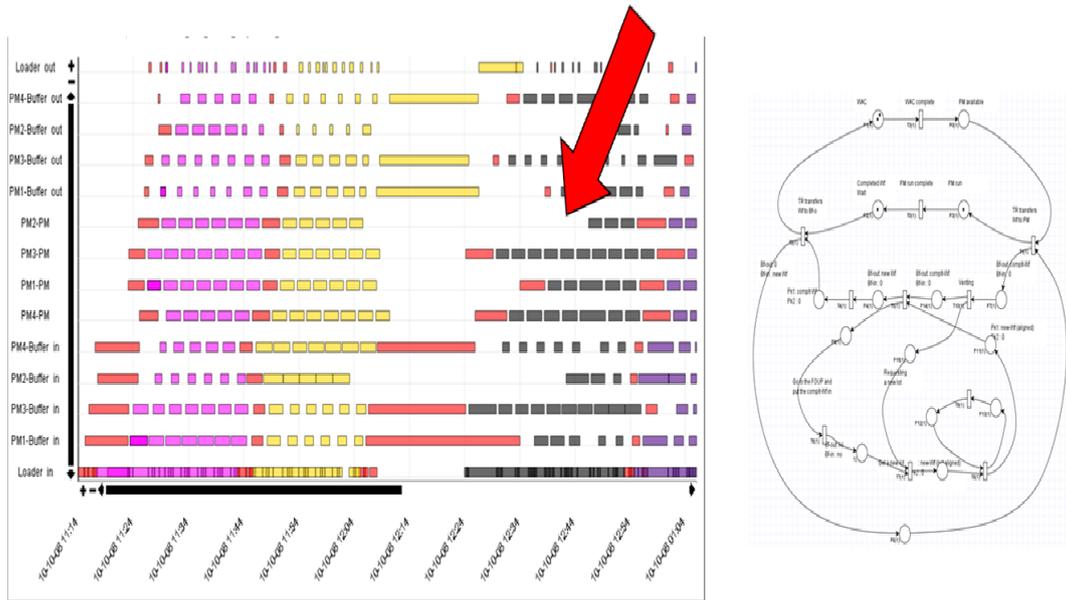
제조 산업에서 센서를 통해 장비의 작업을 실시간으로 파악하는 개념은 이미 제조의 정보화가 급속히 이뤄지던 90년대 MES (Manufacturing Execution System) 개념이 정립될 당시에 소개된 개념이다. 하지만 사물인터넷에서는 이러한 개념을 새롭게 확장하여 해석할 수 있다. 과거의 센서는 센서와 통신하는 (혹은 유선으로 연결하는) 일종의 미들웨어가 센서를 제어하면서 센서는 사전에 정의된

정보를 정해진 방식으로 미들웨어에 전송하는 역할을 하였다. 사전에 모든 것이 정의되어있고 정해진 룰만 따르는 매우 정형화된 장비였다. 하지만 최근 사물인터넷 개념에서는 센서 자체의 지능으로 센서 간의 연동을 통해 미들웨어의 변경이나 세팅 없이 유연하게 정보를 취합할 수 있다. 또한 과거에는 미들웨어에서 정보를 필터링하여 특정 정보만 취합할 수 있었던 반면 최근 센서의 개념은 자체적인 통신망을 통해 다양한 데이터를 자유자재로 송신할 수 있는 개념으로 진화하고 있다.

2.4 빅데이터-사물인터넷 혁신사례

센서 데이터의 특징은 실시간 시그널을 시간에 맵핑해 제공해주는 시간-이벤트 (time-event) 형식을 지니고 있는 것이다. 이러한 시간에 따른 이벤트를 정의해 상황을 설명한 것을 로그 (Log) 혹은 로그 데이터 (Log Data)라 한다.

로그데이터 분석을 통한 실제 사례는 미국 반도체기업 마이크론 테크놀로지 (Micron Technology)사에서 2009년부터 전사적으로 실행하고 있는 반도체 장비 로그데이터 분석 프로젝트에서 찾을 수 있다 [13]. 당시 마이크론 테크놀로지는 각 반도체 장비에장비의 모든 이벤트를 수집하는 로그데이터를 기업의 경쟁력으로 활용 할 수 있는 방안을 모색하며 로그분석 전담 과제를 진행하였다. 약 5개월에 걸친 로그데이터 분석결과는 기존에 알지 못한 새로운 인사이트와 문제점을 파악하는데 성공하였다. <그림 2>는 반도체 장비 내 로봇의 움직임을 로그 데이터로 분석해서 찾은 사례이다. 아래 그림은 로그데이터를 기반으로 로봇의 움직임을 겐트 차트 (Gantt Chart)화 한 것이다. 본 프로젝트에서 겐트 차트 분석에서 특정 이벤트의 오류를 발견할 수 있었다. 챔버 (chamber)에서 작업이 종료되는 다음 작업이 진행될 때 반드시 클리닝작업이 진행되어야 하는데 가끔 이러한 클리닝 작업이 소프트웨어 오류로 작업이 진행되지 않았음을 발견하였다. 비록 1초도 되지 않는 클리닝 작업이고 가끔 발생하는 작업이어서 이러한 오류가 큰 문제점이 없다고 판단 할 수 있다. 하지만 전체 로봇 작업의 순서를 마이닝 기법으로 분석한 결과 이러한 작은 오류는 로봇 전체의 데드락 (deadlock)을 일으키는 단초란 사실을 발견하였다. 만일 로봇이 데드락이 되면 전체 장비가 멈추게 되고 일단 멈춘 장비는 전체 초기화 작업이 진행되어야 다시 작업을 수행 할 수 있다. 즉 아주 작고 사소한 장비의 큰 문제점을 일으키는 문제점의 시작점이란 사실이다.



<그림 2> 로그데이터 분석을 통한 문제점 도출 사례 (붉은 화살표 지점은 장비의 에러 부분 [12])

그렇다면 본 장비를 담당하는 엔지니어 및 관리자는 왜 이러한 문제를 사전에 인지하지 못 했을까? 반도체 장비는 수십여 개의 장비 업체를 공동작업의 결과물이다. 예를 들어 대표적인 반도체 제조업체인 어플라이드 머티리얼사 (Applied Material Advanced Technology - AMAT)사의 PVD장비의 경우 내부 물류반송 로봇은 일본 하네다 (Haneda) 로봇사가 제작하고 이 로봇의 소프트웨어는 미국 내 PRI 솔루션사가 제작한 소프트웨어를 사용한다. 그리고 PRI 솔루션사의 소프트웨어 중 로봇의 로직 부분은 미국 캠브리지 로봇이란 업체가 제작한 로직으로 구동된다. 즉 수 천 개의 하드웨어와 서로 다른 업체에서 개발된 소프트웨어가 결합한다면 어떤 일이 벌어질지를 예측하는 것이 쉽지 않다. 더구나 장비가 반도체 업체에 납품되면 반도체 업체는 자체 생산 프로세스의 생산 효율을 높이기 위해 자체적인 튜닝 작업을 진행한다. 그러나 튜닝이 된 이후 하드웨어와 소프트웨어 로직이 서로 문제없이 연동할 수 있는지는 늘 현장의 고민이기도 하다.

이처럼 현장에서는 반도체 장비가 정해진 로직에 따라 문제없이 하드웨어가 구동될 것이란 것에 기대에 가깝다. 늘 끊임없이 문제점이 도출되고 문제를 해결하는 연속적인 작업이 이뤄지는 이유이기도 하다. 하지만 과거에는 문제가 일어난 후 사후 문제를 해결하는 방식으로 진행되었다면 마이크론사의 경우는 문제를 사전에 파악해서 선제적으로 대응하는 방식을 이러한 로그데이터 분석을 통해 가능하게 했다. 원래 사전에 약속된 로직대로 구현하는 것을 분석하는 것이 아닌 데이터 기반 결과를 바탕으로 결과를 통해서 원래 어떤 의도로 진행되는 것이었을까를 유추하는 방식이 실시간 센서에서 생성되는 로그데이터 분석 통한 방식이라 할 수 있다.

마이크론은 센서 로그데이터의 새로운 가치와 로그데이터 분석 기법을 통해 데이터의 새로운 가능성을 찾았지만 이러한 방식을 전사적으로 수행하는 데는 새로운 문제점에 직면하게 된다. 바로 로그데이터를 수집하는 센서의 독립성이슈다. 반도체 장비의 모든 이벤트를 분석하기 위해서는 로봇의 동작을 인식하는 장비내부에 장착된 센서의 정보뿐만 아닌 장비에 물류를 투입하는 AMHS (automated material handling system - 물류자동반송) 장비의 센서와도 연동이 필요하다. 하지만 장비 내 정보시스템과 AMHS의 장비의 정보시스템은 상호 다른 미들웨어에 의해 작동되어 이들 센서 내 정보교환을 위해서는 미들웨어 교체 및 새로운 대규모 IT 투자가 필요했다. 만일 센서가 미들웨어에 종속되지 않고 자체적인 표준으로 정보를 전송할 수 있는 기능이 있다면 보다 쉽게 통합된 데이터 분석이 가능하다.

현재 대부분 자동화 공장에 설치된 센서는 특정 미들웨어에 종속된 형태로 설계된 센서다. 즉 사전에 각 센서가 어떤 데이터를 수집해서 어떻게 분석될 것인지 계획된 작업만 가능하다. 하지만 시시각각으로 진화하는 제조 환경에서 늘 새로운 데이터와 새로운 분석이 요구되고 이러한 새로운 분석이 요구될 때 마다 새로운 투자를 하는 것은 불가능하다. 이러한 대안으로 제시된 것이 미들웨어를 거치지 않고 바로 센서에서 원천데이터를 직접 인터넷이나 데이터 정보시스템으로 전송하는 기술이 필요한 이유다.

사물인터넷을 통해 제조의 혁신이 가능한 것도 필요하면 쉽게 센서를 부착해서 정보를 취득할 수 있는 편의성이 제공되었기 때문이다. 이른 앞장에서 설명한 스포츠 분석의 혁신과도 일맥상통하다. 기존 센서의 한계점을 사물 인터넷을 통해 극복하고 새로운 인사이트를 도출할 수 있는 기회를 도모할 수 있다.

III. 정리

과거 상상할 수 없었던 다양한 데이터가 쌓여지는 시대에서 빅데이터는 이러한 데이터에서 가치를 창조하는 개념이라 할 수 있다. 즉 특정 기술을 의미하기 보다는 다양한 기술이 융합되어 과거 불가능했던 새로운 혁신을 창조하는 개념이다. 이러한 비즈니스의 혁신과 결합하여 비즈니스 애널리틱스가 탄생하게 되었다. 본 논문에서는 이러한 빅데이터의 개념과 비즈니스 애널리틱스의 배경 그리고 비즈니스 애널리틱스의 핵심인 데이터 기반 의사결정 프레임워크를 소개하였다. 또한 최근 이슈가 되고 있는 사물인터넷과 빅데이터의 관계 및 실제 사례를 소개하였다. 데이터 중심 사회에서 그 개념이 어떤 단어로 정의되건 그것은 중요하지 않다. 단 빅데이터란 개념에서의 주목할 점은 앞으로 더욱더 데이터가 늘어날 것이란 사실과 더욱더 늘어나는 데이터를 분석할 수 있는 기술의 필요성이 더욱더 강조될 것이란 사실이다. 이런 사실만 집중하면 빅데이터의 실체가 무엇인지, 빅데이터 기법이 무엇인지 등 빅데이터를 정의하는 그 자체의 논쟁은 그리 큰 의미를 지니지 못한다.

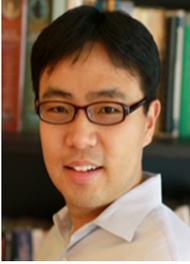
중요한 것은 어느 기업은 이미 쌓이고 있는 데이터를 적극 활용하여 혁신을 이룬 기업이 있고 새로운 다양한 데이터를 기반으로 새로운 학문적 기반을 마련한 연구자들이 있다는 사실에 주목할 필요가 있다.

사물인터넷은 빅데이터의 또 다른 가치를 열고 있다. 센서의 가격하락과 기술발전으로 다양한 센서가 시장에 소개되고 있고 이는 더 많은 데이터가 쉽게 수집될 수 있음을 의미한다. 사물인터넷의 상용화 및 관련 어플리케이션의 등장은 데이터 기반 혁신을 더욱더 가속화 시킬 것이다.

참고문헌

- [1] Ward, Jonathan Stuart, and Adam Barker. "Undefined by data: a survey of big data definitions." arXiv preprint arXiv:1309.5821 (2013)
- [2] Manyika, James, et al. "Big data: The next frontier for innovation, competition, and productivity." (2011)
- [3] Davenport, Thomas H., and Jeanne G. Harris. *Competing on analytics: The new science of winning*. Harvard Business Press, 2007
- [4] McAfee, Andrew, et al. "Big data." *The management revolution. Harvard Bus Rev* 90.10 (2012): 61-67
- [5] Big Data, Analytics Big Data, Analytics and the Path From Insights to Value - *The Magazine - MIT Sloan Management Review*, Winter 2011
- [6] Sathi, Arvind. *Engaging Customers Using Big Data: How Marketing Analytics Are Transforming Business*. Palgrave Macmillan, 2014
- [7] *Too Big to Ignore: The Business Case for Big Data*, Phil Simon, Wiley, 2013 - 당신의 흔적에는 기회가 있다, 장영재/이유진 번역, 한국경제신문사, 2015
- [8] Caro, Felipe, and Jérémie Gallien. "Inventory management of a fast-fashion retail network." *Operations Research* 58.2 (2010): 257-273.
- [9] *코오롱-카이스트 빅데이터와 비즈니스 애널리틱스를 활용한 리테일 연구 - 2014년 연구보고서*, 장영재, 2014
- [10] *경영학콘서트*, 장영재, 비즈니스북스, 2010
- [11] "빅데이터를 활용한 동계스포츠 과학화," 미래부 스포츠 산업융합 연차 보고서, 장영재, 2014
- [12] 화두는 빅데이터 - 결정의 패러다임이 바뀐다, 장영재, 동아비즈니스 리뷰, No, 101, 2011
- [13] Young Jae Jang (2012), *Manufacturing Applications in Big Data*, *The Korean Institute of Communication and Information Science Magazine*, Vol.19, No.11, pp 30-35, September, 2011

장영재(Young Jae Jang)



미국 보스턴대에서 우주 항공공학과를 졸업한 뒤, 미국 MIT 기계공학과에서 석사를 그리고 MIT 슬론 경영대학원에서 경영과학 석사를 받았다. 이어 미국 MIT 공대 기계공학과에서 불확실성을 고려한 생산운영방식에 관해 박사학위를 받았다. 이후 글로벌 반도체기업인 마이크론 테크놀로지 미국 본사에서 기획실 프로젝트 매니저로 과학적 방식을 적용한 원가절감 및 전략적 의사결정을 지원했다.

현재는 카이스트 산업 및 시스템 공학과 교수로 재직 중이며 공급 사슬망 경영, 경영의 수학적 의사결정, 전기 자동차 경제성 분석, 그리고 빅데이터 분석 등 수학 이론과 첨단 기술을 기업 경영의 경쟁력으로 향상시키기 위한 다양한 분야의 연구를 진행하고 있다. 지난 5년간 IEEE 저널 및 관련 학회논문 약 30건을 게재했으며 IEEE Power Electronics 특별호 편집위원 (Associate Editor), IEMS 편집위원 (Associate Editor)를 맡고있다. 또한 International Symposium on Semiconductor Manufacturing Intelligent (ISMI) 2015 공동학회장 (Co-Chair) 및 Asia Pacific Industrial Engineering and Management Systems (APIEMS), 2014 프로그램 위원을 담당하였다. 저서로는 2010년 교보문고 상반기 베스트셀러에 오른 <경영학 콘서트>가 있으며 2011년에는 <경영학 콘서트>를 통해 일반인들에 경영지식을 흥미롭게 전달한 공로로 매일경제신문사의 정진기 언론문화상을 수상하였다.

<Abstract>

Big Data, Business Analytics, and IoT: The Opportunities and Challenges for Business

Young Jae Jang

With the advancement of the Internet/IT technologies and the increased computation power, massive data can be collected, stored, and processed these days. The availability of large databases has brought forth a new era in which companies are hard pressed to find innovative ways to utilize immense amounts of data at their disposal. Indeed, data has opened a new age of business operations and management. There are already many cases of innovative businesses reaping success thanks to scientific decisions based on data analysis and mathematical algorithms. Big Data is a new paradigm in itself. In this article, Big Data is viewed as a new perspective rather than a new technology. This value centric definition of Big Data provides a new insight and opportunities. Moreover, the Business Analytics, which is the framework of creating tangible results in management, is introduced. Then the Internet of Things (IoT), another innovative concept of data collection and networking, is presented and how this new concept can be interpreted with Big Data in terms of the value centric perspective. The challenges and opportunities with these new concepts are also discussed.

Keywords: Big data, Business analytics, IoT

* 이 튜토리얼은 2015년 12월 20일 접수, 편집위원회의 2차례 검토를 거쳐 12월 29일 게재 확정되었습니다.