

KESSIA

Korea Embedded
Software and
System Industry
Association

ISSUE

REPORT

2014. 12

임베디드 SW 산업에 적용 가능한
빅데이터 기술 동향

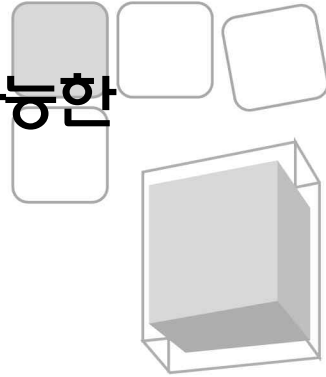
www.kessia.kr

ISSUE REPORT 2014-07

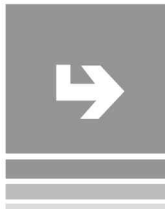


임베디드소프트웨어 · 시스템산업협회
Korea Embedded Software and System Industry Association

임베디드 SW 산업에 적용 가능한 빅데이터 기술 동향



목 차



<요약>	1
1. 빅데이터 기술 이해 및 국내외 동향	3
2. 임베디드 SW 산업 분야에서의 빅데이터 기술 역할	32
3. 임베디드 SW 산업과 빅데이터 기술 전망	54

본 Issue Report에 대한 자세한 문의사항은 아래로 연락주시기 바랍니다.

문의 : 임베디드소프트웨어·시스템산업협회 오정은 선임

T) 02-2132-0755 F) 02-782-1266 E-mail) oje@fkii.org

※ 보고서의 내용을 인용할 때에는 출처를 밝혀 주십시오.

<요약>

최근 IT 기술의 발달에 따라 스마트 폰의 보급과 SNS 활성화, 센싱 기술의 발달과 첨단 연구 장비의 등장으로 과학 기술 뿐만 아니라 사회 전반적으로 빅데이터에 대한 관심이 증폭되고 있다. 다양한 분야에서 빅데이터의 가치가 확인되면서 데이터로부터 차별화된 정보를 추출하여 새로운 과학적 발견이나 기업의 새로운 비즈니스 가치 창출을 활용하는 곳에 빅데이터가 많이 이용된다. 빅데이터는 다양한 종류의 대용량 데이터를 말하며, 데이터를 수집, 저장, 관리, 분석하여 가치 있는 정보를 추출하고 기존 지식과 연계하여 현재 상황을 파악하고 미래 변화를 예측하여 대응 방안을 조언하기 위한 도구와 플랫폼을 포함한 정보화 기술 및 아키텍처를 말한다. 빅데이터의 공통적인 특징은 데이터의 양(Volume), 데이터의 생성 속도(Velocity), 데이터 유형의 다양성(Variety)이며, 이를 3V 특징이라고 한다. 전 세계는 빅데이터의 성장 잠재력과 경제적 가치에 주목하고 있으며, 세계 각국의 정부와 주요 민간 기업들은 빅데이터가 새로운 경제적 가치의 원천이 될 것이라고 기대하고 있다.

전 산업 분야에서 IT 융합화가 빠르게 진행되면서 제품의 지능화, 첨단화가 필수 요소로 되면서 임베디드 SW의 중요성이 급격하게 증가되고 있다. 임베디드 SW는 많은 분야에서 각각의 특성에 맞게 적용되나 국내에서 주력하고 있는 산업 분야는 자동차, 전자, 조선, 기계/로봇, 항공, 의료기기, 에너지 분야로 나눌 수 있다.

본 보고서에서는 임베디드 SW 분야에 적용 가능한 빅데이터 기술 동향에 대해서 살펴보고자 한다. 자동차 산업 분야에서는 구글, 르노, 벤츠, 아우디 등에서 무인자동차를 개발 중으로 임베디드 SW와 빅데이터 기술의 결합을 보이고 있으며, 전자 산업 분야에서는 스마트 홈, 스마트 가전의 사례에 대해 설명한다. 또한 기계/로봇 산업 분야에서는 빅데이터 기반의 인지 컴퓨팅 기술 및 사물통신(M2M)에 대해 설명하며, 항공/의료기기 산업 분야에서는 빅데이터를 활용한 무인항공기, 개인 항공기 개발 기술, 의료기기에 탑재된 임베디드 SW로부터 수집된 데이터들은 빅데이터 분석 기법에 따라 환자의 상태를 파악하고 치료하여 향후, 질병에 대한 예방까지 가능하게 된다. 에너지 산업 분야에서는 스마트 그리드를 활용하여 전력공급자와

소비자가 양방향으로 실시간 정보를 교환하여 에너지 효율을 최적화하고 새로운 부가가치를 창출해내는 차세대 지능형 전력망을 꿈꾼다.

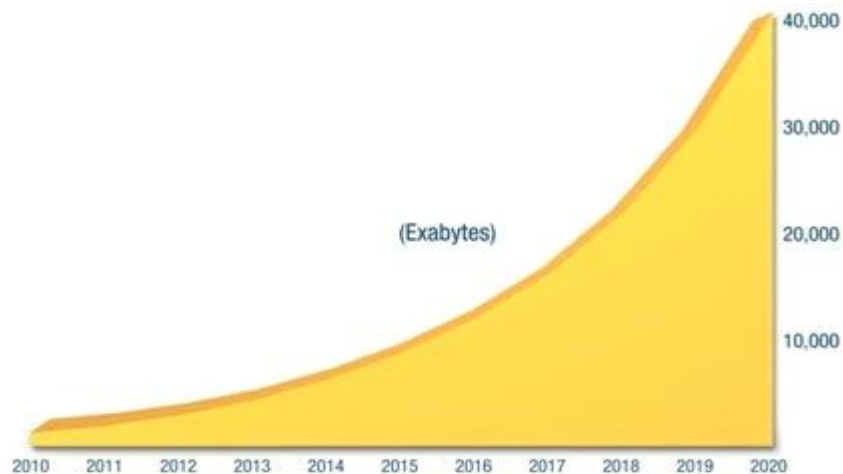
앞으로의 IT 산업 전망에서 가장 흥미로운 트렌드는 사물인터넷(Internet of Things)이다. 사물인터넷은 인터넷으로 서로 연결되어 모든 것들에 대한 정보 생성, 수집되고 공유, 활용되는 것으로 M2M(Machine to Machine), IoT(Internet of Things)를 거쳐 IoE(Internet of Everything)로까지 사용 범위가 확장되고 있다. 사물인터넷의 가장 기본은 IoT를 구동하는 기기들에 탑재된 임베디드 SW이므로 자동차산업, 유틸리티(전기·가스·수도), 의료분야, 가전분야 등의 통신 모듈 및 센서 기업 등에 사업기회가 창출될 것으로 보인다. 현재의 임베디드 SW 산업은 하드웨어 단말에 탑재된 소프트웨어나 기기들이 중심이었으나, 인터넷과 연결하는 사물인터넷 세상이 주류가 되면서 기기들 간의 통신이 가능한 M2M이 핵심이 되고 있다. 또한 기기들 간 주고받는 데이터들이 빅데이터로 수집·처리·분석·예측이 가능하므로 임베디드 SW 산업 분야의 기업들이 빅데이터 기술 동향에 적극적인 관심을 갖고 임베디드 SW 산업에 적용 가능한 빅데이터 기술 확보에 주력해야 할 것으로 전망한다.

1. 빅데이터 기술 이해 및 국내외 동향

1.1 빅데이터 개념

가. 등장배경

- 최근 IT 기술의 발달에 따라 스마트폰, 소셜 네트워킹, 센서 네트워크 등을 통해 등이 일상화되면서 정보의 종류와 양이 급격하게 증가함. 스마트폰은 전 국민이 활용하는 시대를 눈앞에 두고 있고, 소셜 네트워크는 모바일과 만나 갈수록 활용도가 증가하고 있음
- 네트워크 활용도의 증대에 따라 센서와 네트워크와의 결합이 생겨나고 (M2M, IoT, IoE) 다양한 데이터의 결합이 발생함. 각종 전자제품에 센서들이 수개씩 내장되어 있음
- 사물인터넷이 생성하는 디지털 데이터가 폭증함에 따라 2020년에 생성되는 전 세계 디지털 데이터의 양이 44조 기가 바이트(GB)에 달할 것이라고 IDC 조사는 발표함. 이것은 2013년 한해 생성된 디지털 데이터 4.4조 기가 바이트(GB)의 10배에 해당하는 양임



◇ 디지털 데이터의 양은 2년마다 약 두 배씩 증가해 2020년 약 40 제타바이트에 이를 것으로 전망되고 있다. [출처=EMC]

<출처: Digital Universe Report(EMC)>

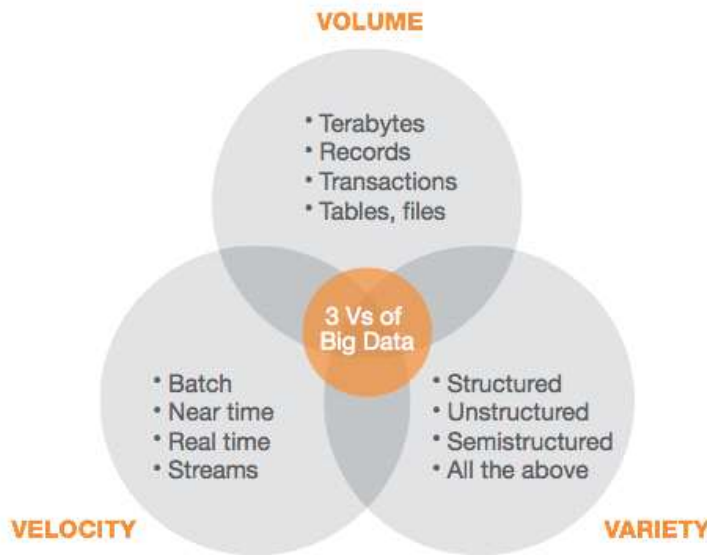
나. 빅데이터 개념

- 맥킨지(McKinsey)는 빅데이터(Big Data)를 일반적인 데이터베이스 체계가 저장, 관리, 분석할 수 있는 범위를 초과하는 규모의 데이터라 정의함
 - 데이터의 양, 생성 주기(실시간 생산), 형식(수치 데이터뿐 아니라 문자와 같은 비정형 데이터 포괄) 등에서 과거 데이터에 비해 규모가 크고, 형태가 다양하여 기존의 방법으로는 수집, 저장, 검색, 분석이 어려운 방대한 크기의 데이터를 의미함
- 가트너(Gartner)는 빅데이터의 3대요소를 데이터의 크기(Volume), 데이터 입출력 속도(Velocity), 데이터의 형태(Variety)로 정의함
 - 데이터의 크기는 스마트 디바이스의 확대·네트워크 고도화로 인해 데이터 량이 증가하면서 제타바이트(Zetabytes) 수준의 데이터를 의미함
 - 데이터 입출력 속도는 데이터 생성-수집-통합-분석-활용단계에서 실시간성 추구를 의미함
 - 데이터의 형태는 기존의 관계형 데이터베이스뿐만 아니라 SNS, 위치 정보, 각종 로그 기록을 비롯해 멀티미디어 등의 비정형 데이터 등 데이터 유형이 다양화되는 것을 의미함
- 빅데이터는 초기에는 수십~수천 테라바이트에 달하는 대용량의 데이터 집합 자체만을 지칭하였으나, 점차 그러한 데이터를 수집, 저장, 관리, 분석하기 위한 기술과 도구, 플랫폼으로 개념이 확대되었으며, 또한 빅데이터를 분석·활용하기 위한 인력과 조직, 빅데이터 산업 등을 포괄하는 용어로 변화되어 사용되고 있음
 - 대용량 데이터를 활용, 분석하여 가치 있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술임 (“빅데이터를 활용한 스마트 정부 구현(안)”, 국가정보화전략위원회, 2011)

- 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고, 데이터의 초고속 수집, 발굴, 분석을 지원하도록 고안된 차세대 기술 및 아키텍처를 뜻함 (IDC, 2011)

다. 빅데이터의 특징

- 빅데이터는 공통적으로 데이터의 양(Volume), 데이터의 생성 속도(Velocity), 데이터 유형의 다양성(Variety) 등 3차원의 특징을 가지고 있으며 정형 데이터를 넘어 문자, 오디오, 비디오, 클릭 스트리밍, 로그 파일 등과 같은 모든 다양한 비정형 데이터를 포함하고 있음

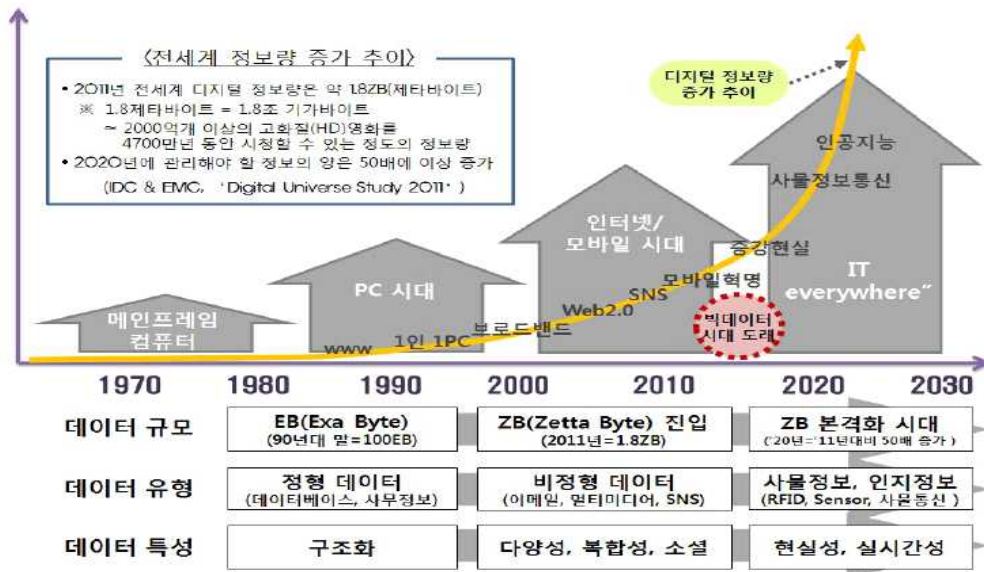


<빅데이터의 특징 - 3V>

- 데이터의 양(Volume): 단순 저장되는 물리적 데이터양의 증가뿐만이 아닌, 이를 관리하고 분석하는 데 어려움이 따르는 네트워크 데이터의 급속한 증가는 빅데이터의 가장 기본적인 특징임
 - 한국정보화진흥원에 의하면 2011년 전 세계 디지털 정보량은 약 1.8

제타바이트로 추정되며, 2020년에는 관리해야 할 정보의 양이 현재의 50배 이상 증가할 것으로 전망함

- Cisco에 의하면 대용량 정보의 공유로 인하여 2010~2015년 사이 전 세계 모바일 데이터 트래픽은 연평균 92%, 인터넷 트래픽은 34% 증가할 것으로 예상함
- 엄청난 크기의 데이터를 수집, 저장, 검색, 공유, 분석 및 가시화를 위한 새로운 저장소 및 이를 관리, 분석하기 위한 기술과 프레임워크가 요구됨



<출처: ICT 발전에 따른 데이터의 변화 방향(빅데이터의 새로운 가능성과 대응전략, 한국정보화진흥원, 2011)>

- 데이터의 생성 속도(Velocity): 실시간으로 빠르게 증가하는 빅데이터는 기존의 정적인 틀을 벗어나 지속적으로 갱신되고 증가하기 때문에 기존의 정보 관리 기법으로 관리할 수 없으며, 이를 효과적으로 저장, 검색, 분석할 수 있는 새로운 기술이 요구됨
- 이처럼 빠르게 생성되며 지속적인 정보의 갱신이 이루어지는 방대한

데이터를 특히 스트리밍 빅데이터로 칭하며, 실시간 처리가 필수적으로 요구됨

- 실시간 처리 외에 장기적 접근 또한 요구됨. 데이터 생산 및 유통, 수집 및 분석 속도의 증가와 이에 대한 실시간 처리와 함께, 장기간에 걸쳐 데이터를 수집·분석하는 장기적 접근이 빅데이터의 두 번째 주요 특징임

□ 데이터 유형의 다양성(Variety): 데이터는 정형화 정도에 따라 정형(structured), 반정형(semi-structured), 비정형(unstructured)으로 분류되며, 과거에는 데이터베이스와 같은 정형 데이터만을 주로 다루어 왔으나, 데이터 생성의 주체와 방식이 다양해짐에 따라 생성된 데이터의 유형 또한 다양해졌으며, 특히 비정형 데이터의 비율이 상대적으로 크게 증가하여, 이를 다루기 위한 기술의 중요성이 커졌다는 것이 빅데이터의 세 번째 주요 특징임

- 정형(structured) 데이터는 고정된 필드에 저장된 데이터를 가리키며, 관계형 데이터베이스 및 스프레드시트 등이 해당되며, 반정형(semi-structured) 데이터는 고정된 필드에 저장되어 있지는 않지만, 메타데이터나 스키마 등을 포함하는 데이터로, XML이나 HTML 텍스트 등이 해당되며, 비정형(unstructured) 데이터는 고정된 필드에 저장되어 있지 않은 데이터로, 언어 처리가 필요한 텍스트 문서와 이미지, 동영상, 음성 등을 포함하는 멀티미디어 데이터 등이 해당됨
- 최근 빅데이터를 이용한 데이터 분석은 고정된 시스템에 저장되어 있지 않은 XML, HTML 등과 같이 데이터베이스 스키마를 포함하는 반정형 데이터를 이용한 분석뿐만 아니라 텍스트, 사진, 오디오, 비디오 형식의 소셜 미디어 데이터나 로그파일 같이 비정형 데이터도 처리할

수 있는 능력을 요구함

- 특히, 빅데이터는 약 80%이상이 비정형 데이터로 이루어져 있으며 다양한 비정형 데이터를 저장, 관리할 수 있는 기술과 비정형 데이터에서 정확한 의미 정보를 추출하고 활용할 수 있는 정밀한 데이터 마이닝 기술 및 추출된 정보를 효과적으로 융합할 수 있는 고급 분석 기술이 요구됨

라. 빅데이터 중요성과 가치

- 전 세계는 빅데이터의 성장 잠재력과 경제적 가치에 주목하고 있고, 세계 각국의 정부와 주요 민간 기업들은 빅데이터가 새로운 경제적 가치의 원천이 될 것이라고 기대함



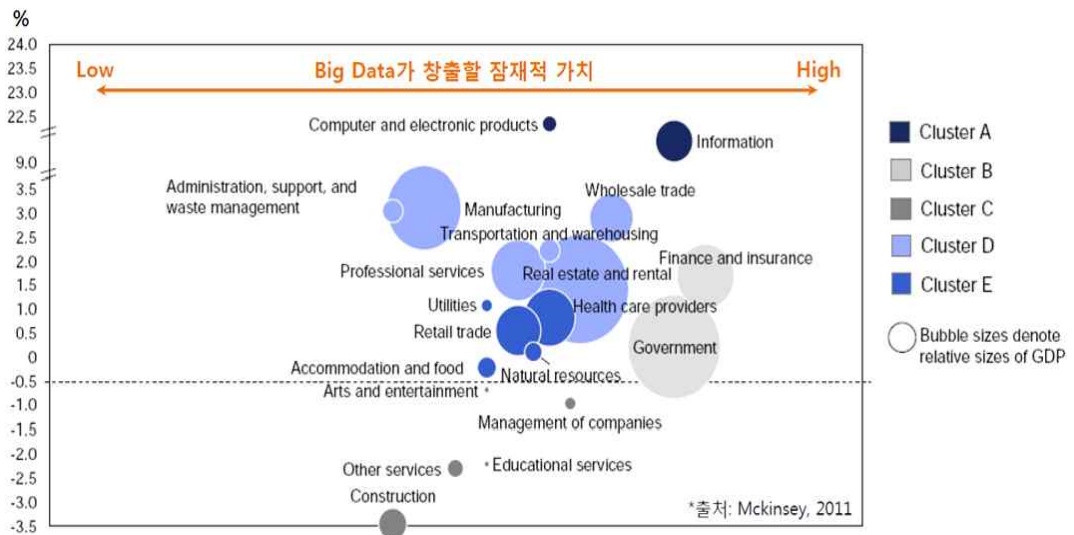
<출처: 빅데이터가 제공하는 경제적 가치(McKinsey(2011))>

- 맥킨지, 가트너 등 빅데이터를 활용한 시장 동향, 신산업 발굴 등 경제적 가치 사례 및 효과를 제시함. 단순히 서비스 차원을 넘어 상품기획, 마케팅 및 R&D의 기반이 되는 지능형 플랫폼으로 성장하여 전 산업의 패러다임을 변화시켜가고 있음

기관명	주요 내용
Economist	<ul style="list-style-type: none"> ▶ 데이터는 자본이나 노동력과 거의 동등한 레벨의 경제적 투입 자본, 비즈니스의 새로운 원자재 역할 ▶ 비즈니스 트렌드 파악, 질병 예방, 범죄 해결 등
Gartner	<ul style="list-style-type: none"> ▶ 데이터는 21세기 원유, 데이터가 미래 경쟁 우위를 좌우 ▶ 빅데이터는 Peak of Inflated Expectations 상태임
McKinsey	<ul style="list-style-type: none"> ▶ 빅데이터는 글로벌 비즈니스 지형을 뒤바꿀 기술 트렌드의 3가지 핵심중 하나임 ▶ 의료, 공공행정 등 5대 분야에서 6천억 불 이상 가치 창출

< 출처: Economist, Gartner, McKinsey의 빅데이터 활용 경제적 가치 >

□ 빅데이터의 활용으로 다양한 분야의 산업에서 경제적 가치를 발생할 것으로 예상함



- Cluster A: 컴퓨터, 전자제품 및 정보통신 분야
 - 빅데이터 활용 가치가 가장 높음
 - 빅데이터 접근 가능 → 혁신 속도 촉진
- Cluster B: 금융, 보험, 정부 분야
 - 장애요인을 극복하면 빅데이터 혜택 가능
 - 고객 특성별 세분화 및 자동 알고리즘 활용
- Cluster C: 건설, 교육서비스, 엔터테인먼트 분야
 - 생산성 증대를 막는 구조적인 장애요인 이 존재하지만, 이를 극복하면 생산성 향상 가능
- Cluster D, E: 부동산, 자연자원, 음식 등
 - 상대적으로 데이터 활용과 접근이 제한적

< 출처: 산업분야에서의 빅데이터 활용(McKinsey,2011) >

- 세계 주요 기관은 빅데이터를 향후 유망 기술로 선정함
 - 세계경제포럼은 2012년 떠오르는 기술로 빅데이터를 선정하고, 가트너(Gartner), IDC(International Data Corporation 등 글로벌 리서치 업체들도 비즈니스 지형을 바꿀 10가지 기술 트렌드 중 하나로 빅데이터를 선정함
 - 우리나라 지식경제부 R&D 전략기획단은 IT 10대 핵심기술 중 하나로 빅데이터를 선정하고 2011년 ‘빅데이터를 활용한 스마트 정부 구현안’을 발표함
- 이제, 빅데이터는 단순히 경제적 가치를 넘어서 국가 경쟁력의 원천이며, 향후 미래사회의 가치를 창출할 엔진으로 인식됨

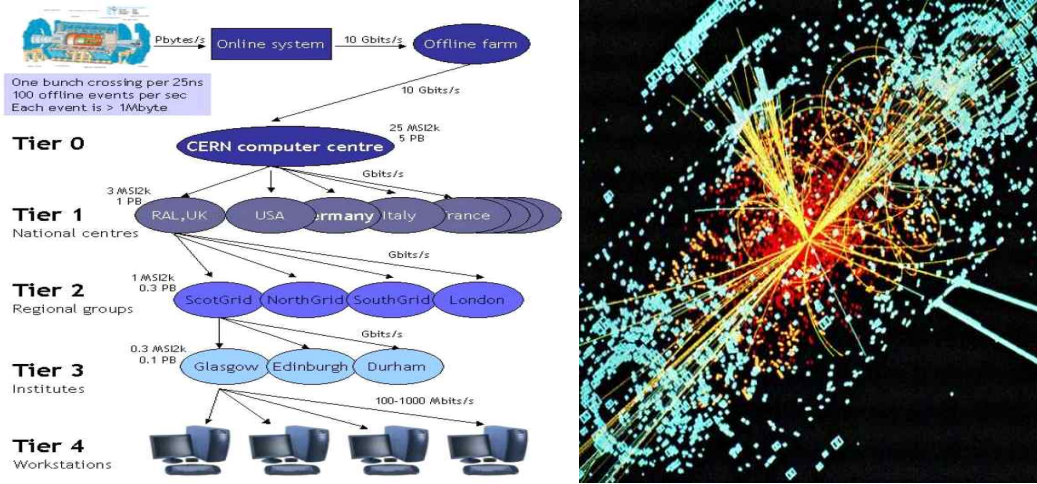
미래사회 특성	빅데이터의 역할
불확실성	통찰력 <ul style="list-style-type: none"> ▸ 사회현상, 현실세계의 데이터를 기반으로 한 패턴 분석과 미래전망 ▸ 여러 가지 가능성에 대한 시나리오 시뮬레이션 ▸ 다각적인 상황이 고려된 통찰력을 제시 ▸ 다수의 시나리오로 상황 변화에 유연하게 대처
리스크	대응력 <ul style="list-style-type: none"> ▸ 환경, 소셜, 모니터링 정보의 패턴 분석을 통한 위험징후, 이상 신호 포착 ▸ 이슈를 사전에 인지·분석하고 빠른 의사결정과 실시간 대응 지원 ▸ 기업과 국가 경영의 투명성 제고 및 낭비요소 절감
스마트	경쟁력 <ul style="list-style-type: none"> ▸ 대규모 데이터 분석을 통한 상황인지, 인공지능 서비스 등 가능(개인화, 지능화 서비스 제공 확대) ▸ 소셜(니즈)분석, 평가, 신용, 평판 분석을 통해 최적의 선택 지원 ▸ 트렌드 변화 분석을 통한 제품 경쟁력 확보
융 합	창조력 <ul style="list-style-type: none"> ▸ 타 분야와의 경합을 통한 새로운 가치 창출 (의료정보, 자동차정보, 건물정보, 환경정보 등) ▸ 인과관계, 상관관계가 복잡한 컨버전스 분야의 데이터 분석으로 안전성 향상, 시행착오 최소화 ▸ 방대한 데이터 활용을 통한 새로운 융합시장 창출

<출처: 신가치 창출 엔진, 빅데이터의 새로운 가능성과 대응 전략(NIA,2011)>

1.2 빅데이터 국내외동향

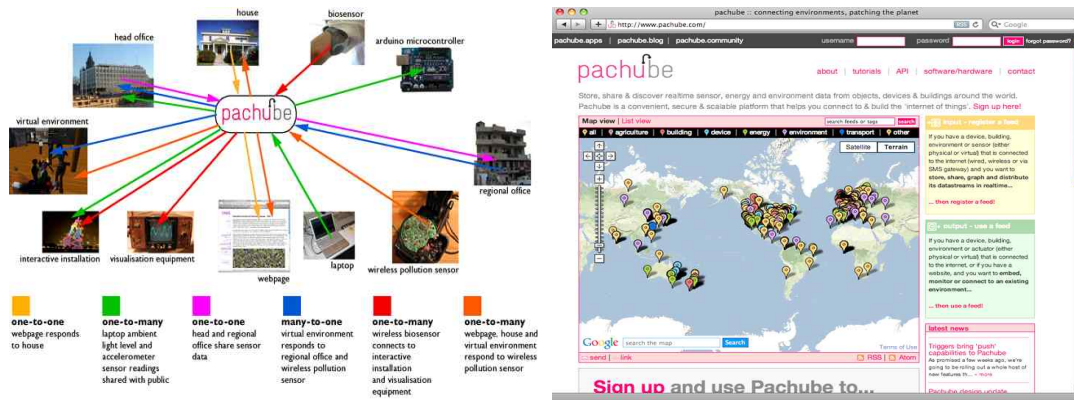
가. 빅데이터 국외 동향

- 유럽 입자물리연구소 : 대형 강입자 충돌기로부터 얻어진 데이터 분석을 통한 물리학 표준모형 완성함
 - 힉스 입자 발견을 위한 대형 강입자 충돌기(LHC) 건설 및 대용량 데이터 분석을 위한 그리드 기술 건설함
 - ▶ 2010년 한 해 동안 300개의 힉스 입자가 만들어졌으며, 수십억 번의 충돌 사건 중 300개의 사건을 찾으면 힉스 입자 발견 가능해짐
 - ▶ 2011년 대형강입자충돌기는 2010년 데이터의 약 150배에 달하는 데이터 1.95 EB, 힉스입자 존재를 확인하기 위해 2012년에는 6 EB 데이터 창출함
 - ▶ 유럽입자물리연구소는 2002년부터 그리드를 전 세계에 걸쳐 건설하였으며 이를 통해 수천 명의 물리학자들이 연구 진행함
 - 새로운 과학적 발견을 위한 대용량 데이터 분석 기술 발전에 기여함
 - ▶ 양성자 충돌을 통해 얻은 대용량 데이터를 축적하고 각 대륙 및 지역별로 분산된 연구자들 간 상호 데이터 교류 방안 지원함
 - ▶ 그리드 기술과 하둡 기술을 이용하여 힉스입자 존재를 확인하기 위해 매해 기하급수적으로 창출되는 데이터에 대한 분석 기술 제공함



<출처: CERN 빅데이터 전송시스템 및 시뮬레이션 예>

- 영국 파큐브 : 공공·민간·개인 등이 다양한 센서를 통해 수집한 대용량 정보 상호 공유 플랫폼 개발함
 - 수많은 센서로부터 입력된 전력, 환경 등의 정보를 개방·공유하는 플랫폼 개발함
 - ▶ 2010년 영국에서 시작되었으며 공공기관, 민간기업, 개인 등이 보유하고 있는 전기, 가전, 휴대폰, 가로등의 센서로부터 제공된 정보를 저장하여 분석·제공함
 - ▶ 공유데이터는 웹 프로그램, 스마트폰 앱 개발 등에 응용 및 활용. 또한, 재난 안전 관리 시스템의 상호 연계 지원함
 - 다양한 센서로부터 얻어진 공공, 민간기업, 개인의 정보를 실시간 분석함으로써 원하는 정보를 빠르게 확보 가능함
 - ▶ 개방형 플랫폼을 통해 누구든지 정보를 가공하여 원하는 애플리케이션 형태로 정보 제공 지원함
 - ▶ 공공기관, 개인, 민간이 가지고 있는 센서 데이터를 통합적으로 분석하여 각종 데이터의 효율적 활용 지원함



<과큐브 개념 및 서비스 예>

- 일본 기상청 : 센서를 통해 수집된 재난현장의 데이터 분석으로 재난대응 능력을 강화함
 - 매년 자연재해와 인위적인 재난으로 인한 피해가 급증함에 따라 체계적인 재난 복구와 인명 구조 인프라 구축의 필요성 증가함
 - ▶ 지진, 해일, 태풍과 같은 자연재해 및 테로 등의 국가적 재난 상황에서 국민의 생명과 재산을 보호하기 위한 시스템 필요함
 - ▶ 4세대 이동 통신 기술, 인공위성망 등을 통하여 화재, 방사능 오염 등 사람이 접근하기 힘든 재난현장의 데이터 수집 가능함
- Apple 시리 : 지능형 음성인식 기술과 위치정보 데이터에 대한 통합 분석을 통해 지능화된 사용자 맞춤형 서비스 제공함
 - 인공지능을 강화한 진보된 음성인식 엔진 개발 및 인공지능 기술을 응용한 사용자 맞춤형 서비스 제공함
 - ▶ 사용자의 위치 정보와 언어를 받아들여 문맥을 이해하고 사용자가 원하는 검색 결과나 기능을 실행해주는 음성인식 기술 개발함
 - ▶ 사용자의 행동 데이터를 계속적으로 분석함으로써 기계 스스로 성능을 향상시켜 사용자가 원하는 최적의 정보 제공함
 - 향후 시리와 다양한 산업 분야의 서비스들과 결합하여 엄청난 IT 산

업으로 성장할 가능성 제공함

- ▶ 음성인식 기술 관련 세계 시장 규모는 2005년 11억 달러에서 2010년 30억 달러로 성장한 데 이어 2013년에는 54억 달러까지 성장할 것으로 전망함
- ▶ 인공지능이 강화된 음성인식 엔진을 기반으로 데이터 분석을 통해 다양한 분야에서의 지능화된 사용자 맞춤형 서비스 솔루션 시장 개척 지원함

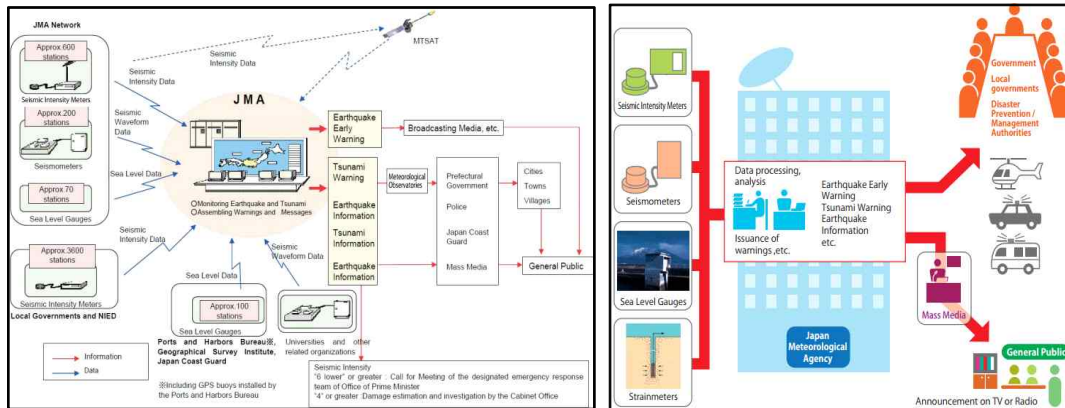


<출처: 시리 홈페이지 및 음성인터페이스 기술의 활용분야>

□ 일본 기상청 : 센서를 통해 수집된 재난현장의 데이터 분석으로 재난대응 능력을 강화함

- 매년 자연재해와 인위적인 재난으로 인한 피해가 급증함에 따라 체계적인 재난 복구와 인명 구조 인프라 구축의 필요성이 증가됨
 - ▶ 지진, 해일, 태풍과 같은 자연재해 및 테로 등의 국가적 재난 상황에서 국민의 생명과 재산을 보호하기 위한 시스템이 필요함
 - ▶ 4세대 이동 통신 기술, 인공위성망 등을 통하여 화재, 방사능 오염 등 사람이 접근하기 힘든 재난현장의 데이터 수집이 가능함
- 다양한 센서 기기를 활용한 신속한 재난 대처로 관련 인력 낭비 감소 및 대응체계 구축에 기여함

- ▶ 기상청이 운영하는 국가지진네트워크 및 국가방재과학연구소에서 구축한 지진네트워크를 통해 지진정보 관측 및 지진현상관측시스템을 통해 분석을 지원함
- ▶ 내각부는 경찰청, 해안보안청, 기상청, 소방청, 방위청 및 지방정부 등과 연계한 긴급 대책본부를 설치하여 피해현황 파악 및 대응체계 구축을 지원함



<출처: 일본 지진정보 모니터링·경보시스템 및 분석과정>

나. 빅데이터 국내 동향

- 한국수자원공사 스마트 워터 그리드 : 센서를 통해 수집된 수도관 정보를 실시간 분석해 누수저감 최적운영 시스템 제공함
 - 최신 IT 기술과 수자원 관리 시스템의 결합을 통해 기존 수자원 생산 및 분배시설의 효율성을 높이는 기술 개발함
 - ▶ 20세기 이후 물 사용량은 6배 증가, 현재 약 11억 명의 인구가 물이 부족한 상황이며, 2025년에는 세계 인구의 약 3분의 2가 물 부족 전망함
 - ▶ 수자원 및 수질관리를 위한 센서 네트워크를 활용하여 연간 1조 3,000궤런의 물 절약과 연간 200만 MWh의 전력 절감 효과를 예상함

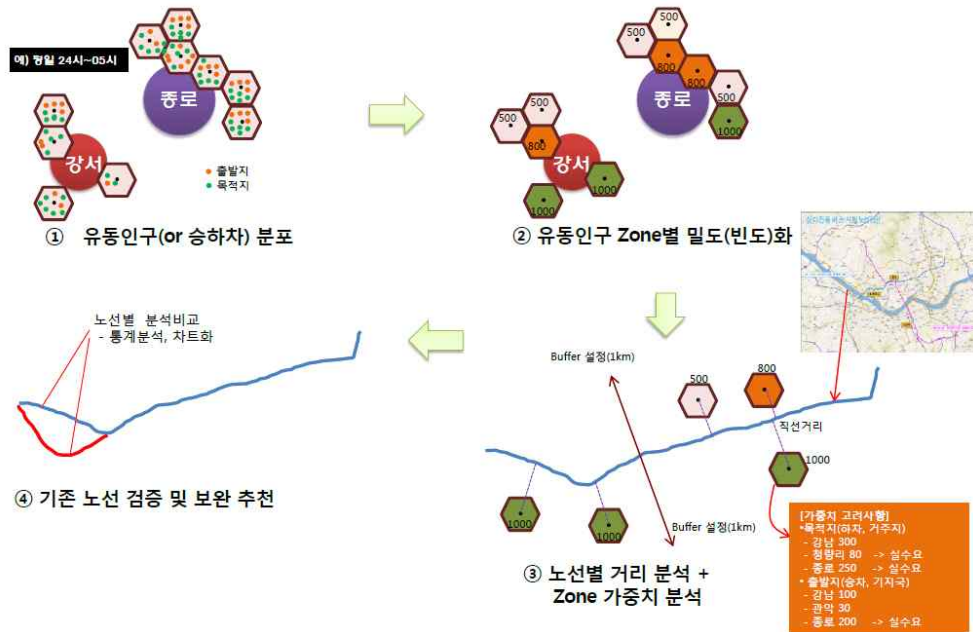
- 효율적인 수자원 관리를 통한 저탄소 녹색성장의 발판 마련함
 - ▶ 룩스 리서치는 2020년까지 163억 원에 달하는 새로운 스마트 워터 그리드 시장이 생겨날 것으로 예측함
 - ▶ 계절별 강우량에 따른 물 부족 현상을 해결할 수 있으며 물 낭비에 대한 사전 대응 지원함



<스마트 워터 그리드 개념도 및 로드맵>

- 서울시 심야버스 노선 최적화: 서울시는 빅데이터를 활용한 심야시간대 심야버스인 ‘올빼미버스’ 노선 최적화를 수립함으로써 시민 생활의 질을 높임
 - 신용카드 결제 데이터, 휴대전화 통화 이력 데이터, 택시 스마트카드 데이터 등을 수집하여 데이터에 의한 정량적 유동인구 분포를 파악함
 - 1.유동인구 밀집도 분석: 일단 서울시를 1km 반경의 1,250개 헥사셀 단위로 구분한다. 그리고 한 달 치 KT 휴대전화 이력 데이터로 심야시간(0시~5시) 통화량 분석을 해서 구역별 유동인구 밀집도를 분석하고 이를 헥사셀 단위로 시각화하는데 성공함
 - 2.유동인구 기반 노선 최적화: 기존 노선의 시간/요일별 패턴을 분석하고, 노선 부근의 유동인구 통계로 가중치를 계산하여, 노선을 최적화함

- 3.유동인구 기반 배차간격 조정: 정류장 단위로 통행량을 추정하고, 통행량을 선의 굵기로 표현하여 hexa셀로 구분된 맵에 시각화함으로써 최종적으로 요일별 배차 간격 조정을 결정함

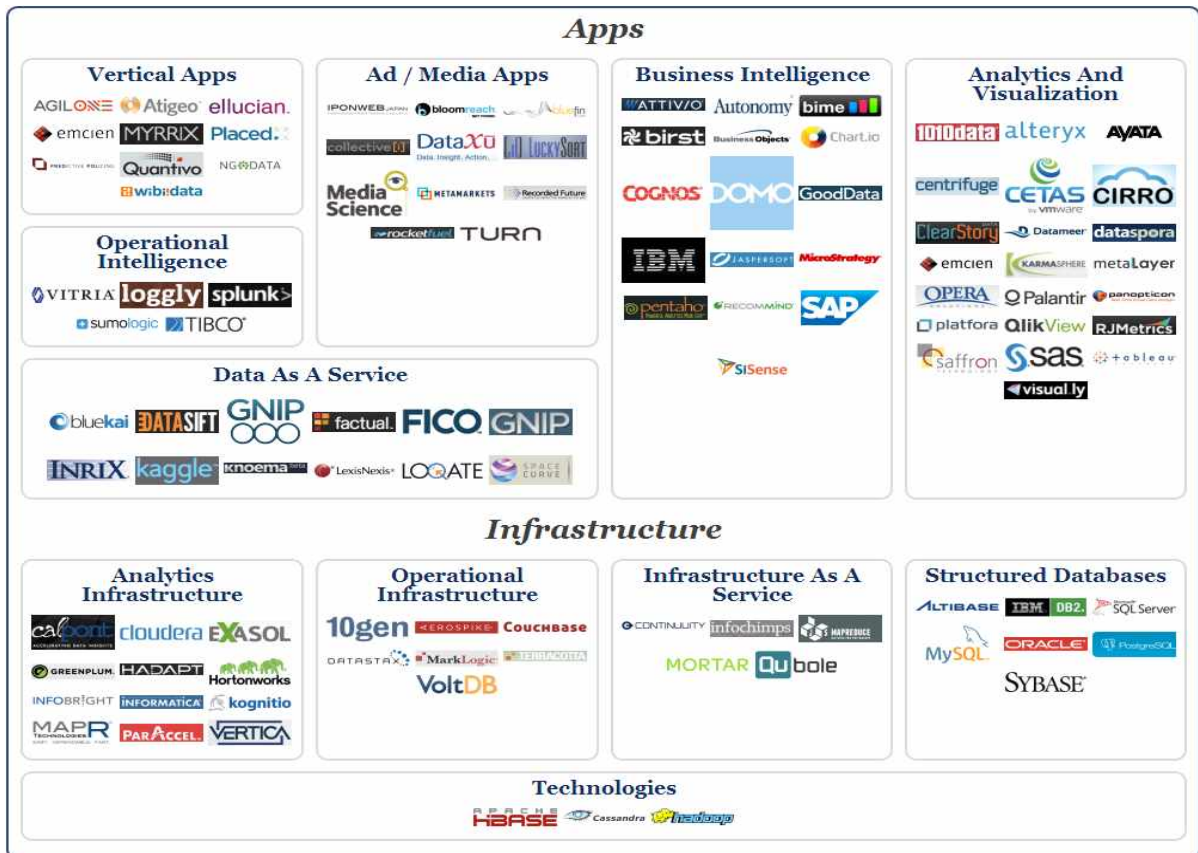


<출처: 심야버스 노선 최적화 분석을 위한 노선 검증 알고리즘>

- 기상청은 빅데이터를 활용하여 전력 수요 예측 오차를 개선해서 연간 1,200억 원의 경제적 효과를 얻음
 - 지난 7년(2004-2011)간의 기상 자료를 분석한 결과 전력 사용량에 미치는 기상요소는 기온(평균, 최고, 최저), 습도, 풍속, 운량, 불쾌지수, 체감온도, 강수량으로 밝힘
 - 빅데이터 분석 기법으로 예측모형에서 과거 특정 기간의 전국 단위 전력 사용량을 예측한 결과, 기온자료만 이용한 경우의 예측 오차가 1.8%인 반면 습도, 운량, 체감온도, 풍속, 강수량 등 기상요소를 추가하였을 때는 예측 오차가 1.3%로 줄어들음을 확인함으로써 약 1,100GWh 정도의 발전량을 줄일 수 있어 약 1,200억 원의 비용을 절감함

1.3 빅데이터 기술

- 소프트웨어의 발전과 더불어 하드웨어와 네트워크 등의 성능 개선과 함께 부품 가격이 싸지면서 기업 내에서는 제품의 지능화 및 타산업과의 융합을 위한 핵심기술로 빅데이터 솔루션 및 기술들에 대한 관심이 점점 높아지고 있음



<출처: Big Data Landscape(<http://www.bigdatalandscape.com/>)>

- 빅데이터로부터 유용한 지식을 활용하기 위해서는 관련된 데이터를 수집하고, 수집한 데이터로부터 불필요한 데이터를 필터링하거나 가공하는 단계, 이러한 데이터를 저장하고 관리하면서 유용한 지식이나 데이터 간의 관계분석을 통해 의미 있는 지식을 발굴, 가치를 발견하는 등의 분석

처리 과정을 거치게 됨

요소 기술	설명	해당 기술
빅데이터 수집	조직내부와 외부의 분산된 여러 데이터 소스로부터 필요로 하는 데이터를 검색하여 수동 또는 자동으로 수집하는 과정과 관련된 기술로 단순 데이터 확보가 아닌 검색/수집/변환을 통해 정제된 데이터를 확보하는 기술	ETL/크롤링 엔진/로그 수집기/센싱/RSS, Open API 등
빅데이터 공유	서로 다른 시스템간의 데이터 공유	멀티 테넌트 데이터 공유/협업 필터링 등
빅데이터 저장	작은 데이터라도 모두 저장하여 실시간으로 저렴하게 데이터를 처리하고, 처리된 데이터를 더 빠르고 쉽게 분석하여, 이를 비즈니스 의사 결정에 바로 이용하는 기술	병렬 DBMS/하둠(Hadoop)/NoSQL 등
빅데이터 처리	엄청난 양의 데이터의 저장·수집·관리·유통·분석을 처리하는 일련의 기술	실시간 처리/분산 병렬 처리/인-메모리 처리/인-데이터베이스 처리
빅데이터 분석	데이터를 효율적으로 정확하게 분석하여 비즈니스 등의 영역에 적용하기 위한 기술로 이미 여러 영역에서 활용해 온 기술임	통계 분석/데이터 마이닝/텍스트 마이닝/예측 분석/최적화/평판 분석/소셜 네트워크 분석 등
빅데이터 시각화	자료를 시각적으로 묘사하는 학문으로 빅데이터는 기존의 단순 선형적 구조의 방식으로 표현하기 힘들기 때문에 빅데이터 시각화 기술이 필수적임	시간시각화/분포시각화/관계시각화/비교시각화/공간시각화/인포그래픽

<출처: 빅데이터 활용단계에 따른 요소기술별 추진동향과 시사점>

가. 빅데이터 수집 기술

- 빅데이터 수집 기술은 조직내부와 외부의 분산된 여러 데이터 소스로부터 필요로 하는 데이터를 검색하여 수동 또는 자동으로 수집하는 과정과 관련된 기술로 단순 데이터 확보가 아닌 검색, 수집, 변환을 통해 정제된 데이터를 확보하는 기술을 의미함 (출처: 빅데이터 활용단계에 따른 요소기술별 추진동향과 시사점)



<출처: “Big Data Capture”, http://www.citsoft.net/?page_id=523>

- 스크라이브(Scribe), 척와(chukwa), 플룸(flume) 등의 다양한 데이터 수집 기술을 통해 수집함
- 수집된 데이터를 변환, 불필요한 데이터를 필터링하는 과정까지 포함함
- 데이터의 구조화된 정도에 따라 정형(Structured), 반정형(Semi-Structured), 비정형(Unstructured)으로 구분하고 있고, 최근은 산업 분야에서 발생한 비정형 스트림 데이터에 대한 처리에 대한 관심이 높아짐. 정형 데이터는 로그 수집기를 통해 수집하며 조직 외부에 존재하는 비정형 데이터는 크롤링, RSS Reader, 또는 소셜 네트워크 서비스에서 제공하는 Open API를 이용한 프로그래밍을 통해 수집함
- 데이터 가공·분석 전문 기업들은 기존의 정형 데이터만을 다루던 틀에서 벗어나 비정형 데이터가 주류를 이루는 빅데이터로 확장하기 위해, 정형/비정형 빅데이터를 통합 운용할 수 있는 기술 및 제품 개발로 전환하고 있음
- 전통적인 관계형 DBMS 기술은 비정형 빅데이터를 분석하는 데 한계가 있고, 비정형 빅데이터에서 중요한 정보를 발굴하고자 하는 요구가 증가하면서 기존의 데이터웨어하우스(DW) 업체들이 정형/비정형 빅데이터

통합 분석 체제로의 전환을 시도하고 있음

- 데이터웨어하우스(DW) 업체인 EMC는 2010년에 그린플럼을 인수한 뒤 최근 비정형 데이터 저장을 위해 하둡을 탑재한 장비를 출시했으며, 뒤이어 2011년에는 무공유 MPP 기반 관계형 데이터베이스(RDBMS)와 엔터프라이즈급의 아파치 하둡을 통합시켜 비정형 데이터와 정형 데이터를 상호 연계해 처리할 수 있는 빅데이터 분석 플랫폼 ‘EMC 그린플럼 DCA’를 출시함
- 테라데이타(Teradata Corp.)는 2011년 아스터 데이터 시스템을 인수하여 대용량 데이터를 병렬처리하고 이를 데이터베이스에서 분석하며 그래프로 분석할 수 있는 기술을 보유하게 됨. 컬럼과 행 기반 분석 기술을 하이브리드 형태로 제공하는 신제품 ‘테라데이타 데이터베이스 14’를 통해, 데이터의 병목현상을 줄이고 데이터베이스 자체에서 워크로드 관리와 보안, 데이터 마이그레이션, 압축 등 인텔리전스 자동화를 제공해 성능을 향상시킴
- SAP는 2011년 인메모리 컴퓨팅으로 실시간 분석이 가능하고 비즈니스 및 소셜 데이터의 정형/비정형 데이터를 통합하는 ‘비즈니스 오브젝트 BI/EIM 4.0’을 발표함
- SAS는 2011년 비즈니스 거래, 소셜 미디어, 위치 정보, 멀티미디어를 포함한 방대한 양의 정형/비정형 빅데이터의 활용 가치를 극대화하기 위한 빅 애널리틱스 전략이 반영된 ‘SAS 9.3’을 출시함

나. 빅데이터 저장 기술

- 빅데이터 저장 기술은 작은 데이터라도 모두 저장하여 실시간으로 저렴하게 데이터를 처리하고, 처리된 데이터를 더 빠르고 쉽게 분석하도록 하여 비즈니스 의사 결정에 바로 이용하도록 만드는 기술임 (출처: 빅데

이터 활용단계에 따른 요소기술별 추진동향과 시사점)

- 대용량 데이터의 저장과 관리, 분석 기술은 분산·병렬 기반의 클라우드 컴퓨팅 기술과 결합하여 하둡을 중심으로 한 빅데이터 분석 에코시스템으로 통합되고 있으며, 특히, 비정형 데이터를 다루기 위한 분산·병렬 기반의 지식 추출 및 인공 지능 기술이 분석의 핵심 기반으로 부상함
- 구글이나 애플, 야후 등에 의해 요소기술로서 상당한 완성도에 도달했으며 오픈 소스로 Hadoop의 HDFS/Hbase, Cassandra, MongoDB 등이 대표적임. 이 중 하둡(Hadoop)은 저렴한 비용으로 빅데이터 시스템을 구축할 수 있는 장점 때문에 현재 빅데이터의 대표적인 기술로 자리 잡음
- 이외에도 SAN(Storage Area Network), NAS(Network Attached Storage)과 같이 기존저장 기술도 있고, Amazon S3나 OpenStack Swift와 같은 클라우드 파일 저장 시스템, GFS(Google File System), HDFS(Hadoop Distributed File System)와 같은 분산 파일시스템 등이 모두 대량의 데이터를 저장하기 위한 기술임
- 분산·병렬 DBMS와 Map-Reduce 방식은 대용량의 데이터를 분산된 데이터 저장소에 저장하고, 복잡한 태스크를 클러스터링된 다중 컴퓨팅 파워를 병렬로 처리함으로써, 계산 복잡도가 매우 높거나 많은 자원을 활용해야 하는 빅데이터 분석의 문제를 비교적 쉽게 해결함
- 기존의 RDBMS 기술은 하나의 시스템이 모든 영역에 맞춰 사용될 수 있도록 만들어져 왔으나, 빅데이터 기반의 OLAP(Online Analytical Processing), 텍스트 처리, 스트리밍 처리, 고차원 데이터 처리 등 지식 추출 기술을 위해서는 특화된 아키텍처를 가진 시스템이 요구됨
- 분산·병렬 DBMS는 전통적인 RDBMS에서 발전한 형태이며, MPP 구

조를 취하고 있는 경우가 많음. 또한 대부분 분산·병렬 DBMS를 개발하는 기업이나 단체는 IT 대기업에 인수되어 부속 모듈(appliance) 형태로 발전하고 있으며, 대표적인 분산·병렬 DBMS에는 VoltDB, SAP HANA, Vertica, Greenplum, IBM Netezza data warehouse 등이 있음

□ 하둡(Hadoop)을 중심으로 대용량 데이터 분석에 필요한 다양한 기술들이 통합되고 응용되면서 하나의 에코시스템을 이루어가고 있고 관련 솔루션 업체들이 등장하고 있음

- 아파치 마하웃(Apache Mahout) 프로젝트는 다양한 중요한 마이닝 알고리즘들을 하둡 프레임워크상에서 구현해서 오픈소스로 제공함. 현재 많은 사람들이 마하웃의 알고리즘을 최적화해서 자신들의 각 분야에서 활용에 따라 하둡 기반의 대용량 마이닝 알고리즘을 제공하는 주요 소스가 될 것으로 예상함
- 미국의 MapR은 Map-Reduce의 소스코드를 수정해서 실시간 데이터를 처리할 수 있도록 하거나 아예 하둡의 소스코드를 수정해 현재 하둡이 가지고 있는 네임노드 가용성 문제, HDFS와 POSIX와의 연계 문제 등을 개선하여 사용 버전을 만들어 사업을 시작함

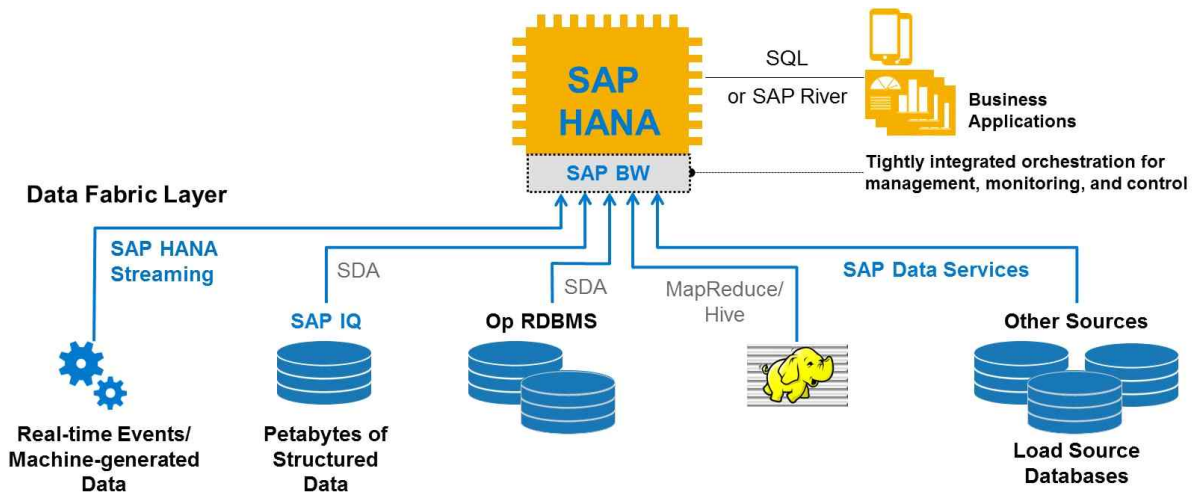
다. 빅데이터 처리 기술

□ 빅데이터 처리기술은 엄청난 양의 데이터를 저장·수집·관리·유통·분석을 처리하는 일련의 기술로써 수많은 사용자 요청을 실시간으로 처리한 후 처리 결과를 반환하는 기술임

- 최근 빅데이터 분석에서 ‘처리 속도’가 중요해지면서 DB가 아니라 메모리에 얹어 실시간으로 활용, 분석하는 인메모리(In Memory) 기술이 주목받고 있음. 인메모리 기술은 단지 분석 솔루션에서 실시간 쿼리를 처리하는 데만 사용되는 것이 아니라 신속하게 데이터를 처리하

기 위한 모든 부분에서 활용 가능성이 거론되고 있음

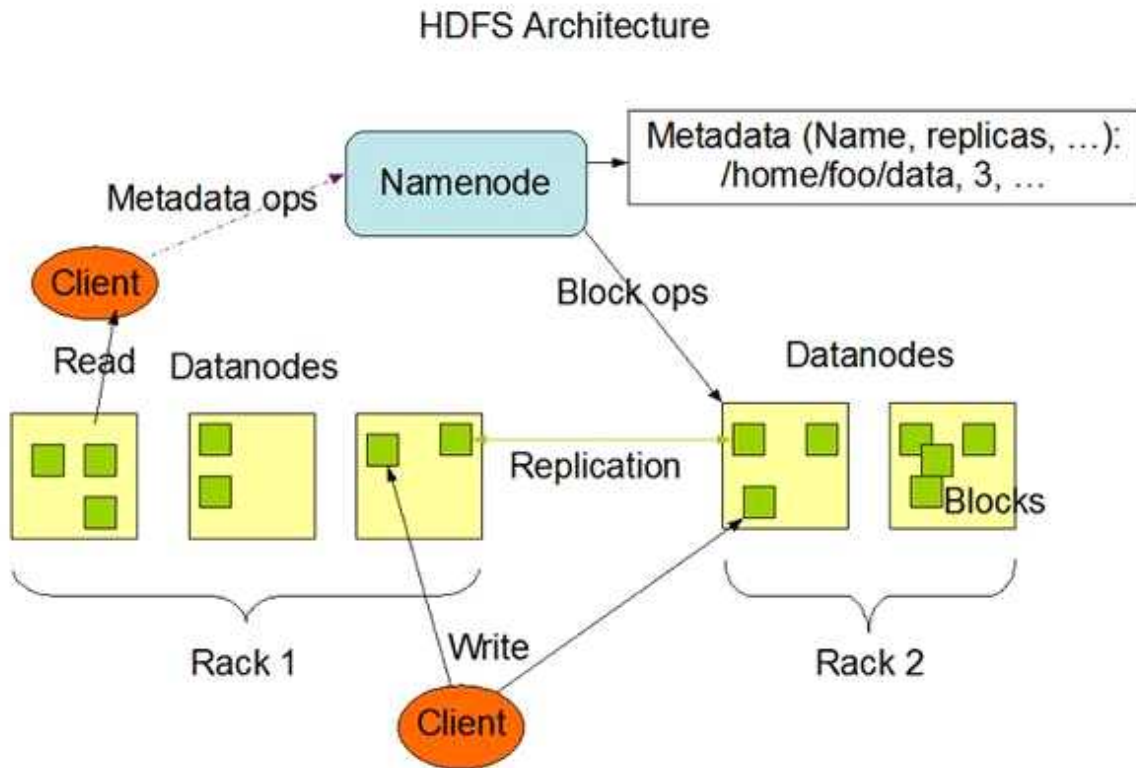
- SAP가 선보인 HANA가 대표적인 인메모리 기반 빅데이터 솔루션으로 인메모리 플랫폼인 HANA를 중심에 두고 비즈니스를 전개하고 있음. SAP HANA는 인메모리DB를 기반으로 방대한 양의 데이터를 실시간으로 검색하고 분석할 수 있는 기능을 제공하면서 빅데이터 수집, 분석 플랫폼임



<출처: SAP BW as part of SAP's in-memory data fabric architecture>

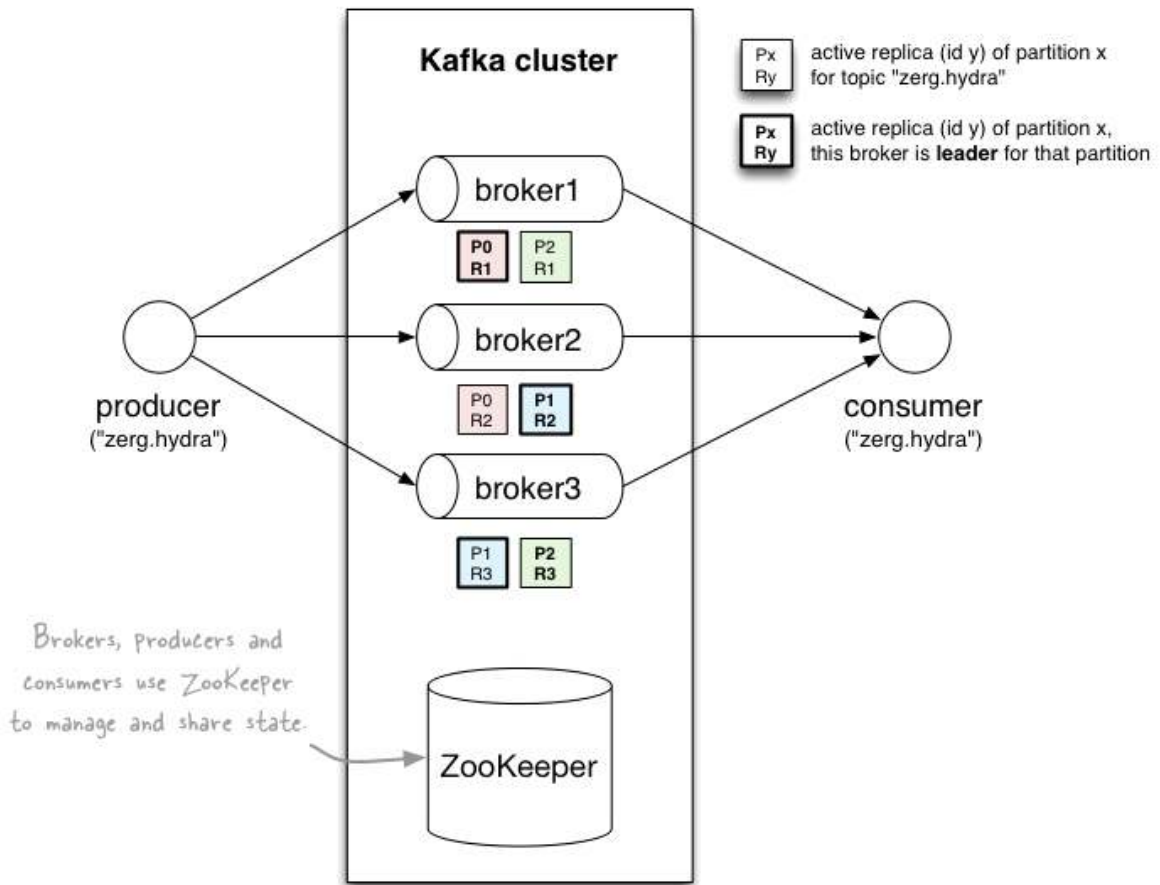
- 빅데이터 처리는 분산처리 시스템인 하둡과 MongoDB, Kafka, 병렬 처리 모델인 맵리듀스, 인메모리 기술인 에스퍼와 레디스 등으로 발전하고 있음
 - 분산처리 기술 중 오픈소스인 아파치 하둡(Hadoop)은 구글의 GFS (Google File System)를 대체할 수 있도록 분산 파일시스템(HDFS)과 MapReduce를 구현한 빅데이터 처리 기술의 대표적인 프레임워크로 하둡의 분산 파일시스템은 여러 기계에 대용량 파일들을 나눠서 저장하고, 데이터들을 여러 서버에 중복해서 저장을 함으로써

데이터 안정성을 확보함



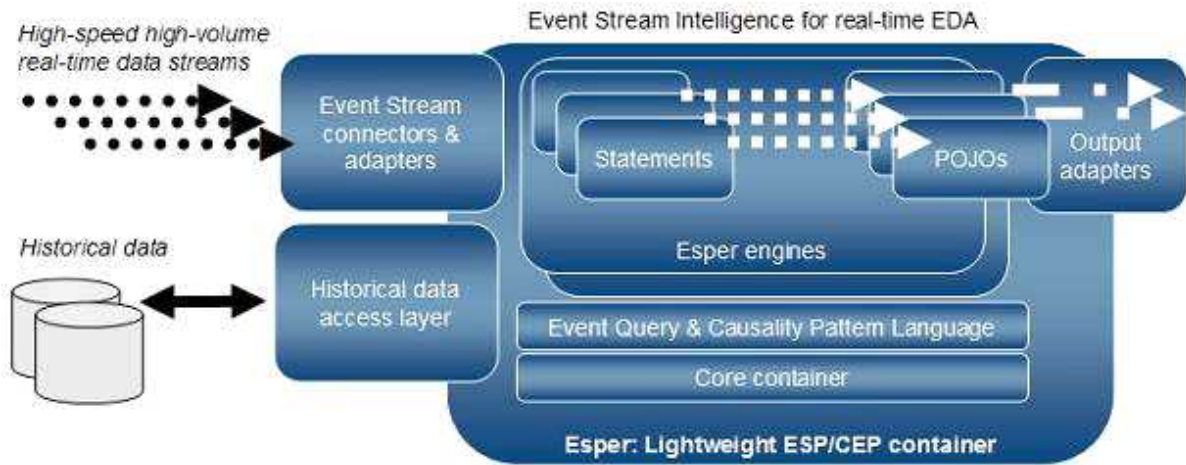
<출처: 하둡 분산 파일 시스템 처리 구조>

- Kafka는 링크드인에서 공개한 대용량 실시간 처리를 위해 사용하는 메시징 시스템으로 Pub-Sub 구조로 되어 있음. 단독으로 실시간 처리를 구성하지 않고 스톰(Storm), 하둡(Hadoop), HBase 등과 연동해서 활용하고 있음. 링크드인(LinkedIn), 트위터(Twitter), 넷플릭스(Netflix), 포스퀘어(Foursquare) 등 대용량을 다루는 업체들이 주로 카프카를 사용하고 있음



<카프카 서버 구성>

- 에스퍼(Esper)는 실시간 처리를 위한 인-메모리 기술 중 CEP (Complex Event Processing) 처리 기술로 CEP란 여러 이벤트 소스로부터 발생한 이벤트를 대상으로 의미 있는 데이터를 추출하여 대응되는 액션을 수행하는 오픈소스 기술. 실시간 트리밍 데이터를 처리하는 이벤트 기반 아키텍처로 구성되어 있으며 특정 이벤트만 필터링해서 필요한 작업을 수행함



<출처: Esper의 이벤트 흐름 아키텍처>

- ▶ 레디스 (Redis)는 인메모리 기술로서 키/값 저장소로 구성되어 있어 NoSQL 데이터베이스로 분류함. 메시지 큐(Message Queue), 공유 메모리(Shared Memory) 용도로 사용될 수 있어 인스타그램, Stack Overflow, Digg와 같은 기업들이 실시간 데이터 스트리밍 처리를 위해 활용되고 있음. 레디스는 인메모리 DB라 빠른 속도가 장점이지만 큰 용량의 데이터를 담기엔 공간 제약이 크므로 실시간 처리는 인메모리에서, 보관은 디스크 기반 스토리지로 하는 구조가 성능과 효율을 함께 달성할 수 있음

라. 빅데이터 분석 기술

- 기업 임원들이 자신들의 본능적인 비즈니스 감각에 기댄 의사 결정이 아닌 빅데이터를 활용한 데이터 분석 결과를 바탕으로 주요 의사 결정을 내리는 것으로 조사됨에 따라 빅데이터 분석에 대한 중요성이 강조되고 있음
- 빅데이터 분석 기술과 방법들은 기존 통계학과 전산학에서 사용되던 데이터 마이닝, 최적화, 소셜 네트워크 분석 등이 주로 활용됨

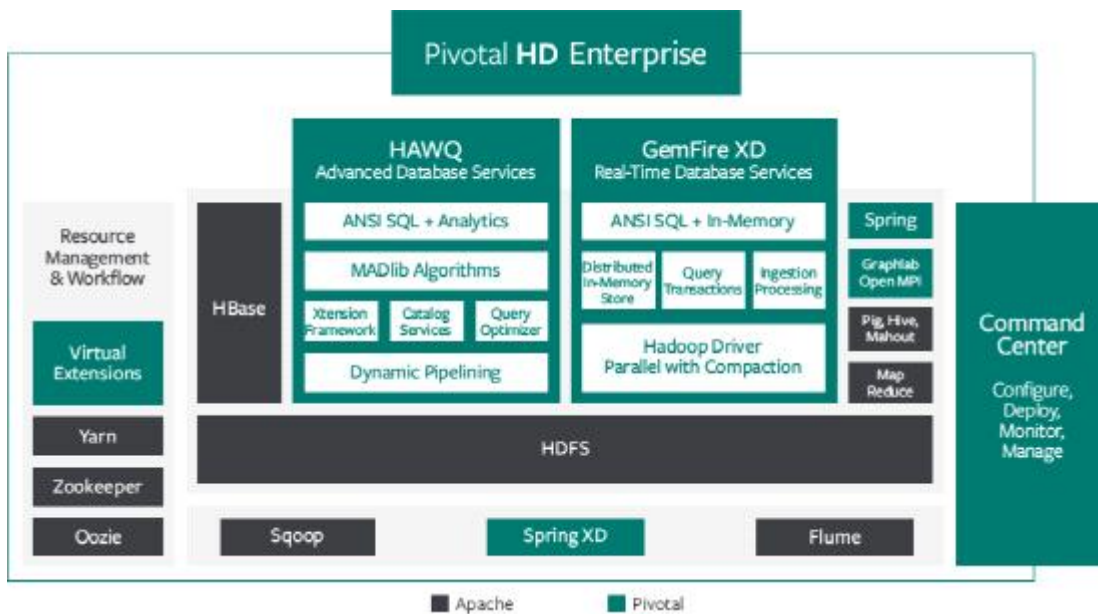
- 빅데이터 분석은 구조화 수준이 낮고 매우 방대한 데이터를 다루므로 더 이상 전통적인 데이터 처리 기술과 분석 기법으로는 사용자의 요구사항을 만족시킬 수 없어, 최근 대용량의 비정형 데이터를 정형화하는 기술인 마이닝 기술이 핵심 기술로 부상하고 있음
 - 데이터 마이닝(data mining)은 대규모로 저장된 데이터 안에서 체계적이고 자동적으로 통계적 규칙이나 패턴을 찾아내는 것임
 - 주요 기법으로 OLAP(On-Line Analytical Processing), 군집 분석(Cluster Analytics), 연결 분석(Link Analytics), 사례 기반 추론(Case-Based Reasoning), 연관성 규칙 발견(Association Rule Discovery), 인공 신경망(Artificial Neural Network), 의사 결정 나무(Decision Tree), 유전자 알고리즘(Genetic Algorithm) 등이 있음
 - 텍스트 마이닝은 구조화되지 않은 대규모의 텍스트 집합으로부터 새로운 지식을 발견하는 과정으로 텍스트 문서 전처리 및 패턴 분석 등의 단계를 가지며, 순환 구조로써 지속적인 피드백을 수행하는데 아래와 같은 분야에 적용하여 결과를 도출할 수 있음
 - ▶ 분류(Classification): 일정한 집단에 대한 특정 정의를 통해 분류 및 구분을 추론 (예: 경쟁자에게로 이탈한 고객)
 - ▶ 군집화(Clustering): 구체적인 특성을 공유하는 군집을 발견. 군집화는 미리 정의된 특성에 대한 정보를 가지지 않는다는 점에서 분류와 다름 (예 : 유사 행동 집단의 구분)
 - ▶ 연관성(Association): 동시에 발생한 사건간의 관계를 정의함 (예: 장바구니안의 동시에 들어가는 상품들의 관계 규명)
 - ▶ 연속성(Sequencing): 특정 기간에 걸쳐 발생하는 관계를 규명함. 기간의 특성을 제외하면 연관성 분석과 유사함 (예: 슈퍼마켓과 금융 상품 사용에 대한 반복 방문)

- ▶ 예측(Forecasting): 대용량 데이터집합내의 패턴을 기반으로 미래를 예측함 (예: 수요예측)
- 현실 마이닝(Reality Mining)은 2008년 미래 유망 기술로 MIT Media Lab에서 발표한 개념으로 사람들의 행동 패턴을 예측하기 위해 사회적 행동과 관련된 정보를 기기(휴대폰, GPS 등)를 통해 얻고 분석하는 기법으로 스마트폰 등 모바일 기기들을 통해 현실에서 발생하는 정보를 기반으로 연관관계와 행동 양태 등을 추론함
- 지식 추출 기술은 현재 단순한 정보의 검색을 벗어나 70%이상의 비중을 차지하는 비정형 데이터를 대상으로 자연어 처리 기법을 통해 구조적이고 유의미한 개체 및 관계를 추출하는 연구로 발전하는 과정에 있으며 계속적으로 그 중요성을 인정받고 있음
 - 지식추출 기술은 비정형 데이터의 홍수 속에서 많은 관심을 받고 있는데, 데이터 마이닝 커뮤니티 사이트인 KDnuggets에서 수행한 2008년도 설문 결과, 55% 이상의 사용자가 그들의 프로젝트에서 지식추출 기법을 사용한다고 답했고, 2009년에는 2008년과 같거나(42%) 많이(52%) 사용할 것이라는 대답이 94%를 차지함
 - 지식추출의 응용인 엔터프라이즈 검색 S/W의 시장 규모는 2010년에 1.2억 달러를 초과할 것으로 예상되고(Gartner, 2008), 2006년 지식추출 S/W 시장 규모는 50-100백만 달러이고 매년 40-60%의 성장률을 보임(Monash, 2006)
- 예측 분석은 과거 자료와 변수 간의 관계를 이용하여 관심이 되는 변수를 추정하는 것으로써 앞서 언급된 통계 분석, 데이터 마이닝 및 텍스트 마이닝 기술들을 기반으로 예측 분석을 수행하게 됨
- 최적화는 주어진 가능한 결과들에 대한 평가를 수행하여 최적의 결과를 도출하는 것으로 비즈니스 환경에서 취할 수 있는 여러 가지 대안들 중

제시된 전략을 평가하고 최적의 대안을 선택하도록 도와줄 수 있는 필수적인 분석 기술임

- 소셜 네트워크 분석은 소셜 네트워크 연결 구조 및 연결 강도 등을 바탕으로 사용자의 명성 및 영향력을 측정하는 기술로써, 수학의 그래프 이론에 뿌리를 두고 있어 주로 마케팅을 위하여 소셜 네트워크상에서 입소문의 중심이나 허브 역할을 하는 사용자를 찾는 데 주로 활용됨
- 실시간 분석, 준 실시간 분석 등과 같은 정확성 보다는 분석 속도에 초점을 두는 분석 기술로 분석에 필요한 모든 가용한 데이터를 활용하여 사용자의 요청에 대한 분석을 수행하고 빠르고 적시에 지식(분석 결과)을 제공해 줄 수 있는 분석기법과 실시간 분석 등을 위하여 인-데이터베이스 분석, 인-메모리 분석, 다중 프로세스를 활용하는 MPP 등과 같이 보다 빠른 지원 기술을 이용할 수 있음
 - 사물인터넷(IoT)의 성장과 함께 다양한 디바이스에서 발생한 데이터들(센서데이터)은 포맷이 다양할 뿐만 아니라 스트림 형태로 데이터가 실시간으로 출력됨에 따라 실시간 분석 기술이 중요해지고 있음
 - 빅데이터의 3V 특성을 모두 포함하는 데이터들을 처리하기 위해서 하둡 기반으로 데이터를 저장, 파싱, 변환, 가공 및 분석을 하는 아키텍처를 일반적으로 도입하고 있음
 - 하둡의 맵리듀스 처리는 일괄 배치 처리하는 특성을 가지고 있어 실시간 분석요건에는 적합하지 않으므로 인메모리기반의 Spark를 활용할 수 있음
 - 하지만, 실제 방대한 데이터 가운데 의미 있는 데이터는 1%라 가정할 경우 이러한 데이터 전체를 서버에서 분석하기에는 불필요한 데이터 처리 비용 및 네트워크 입출력 시간 등이 많이 소요되므로 효율적으로 서버를 활용하기 위한 데이터 필터에 대한 요구도 많음

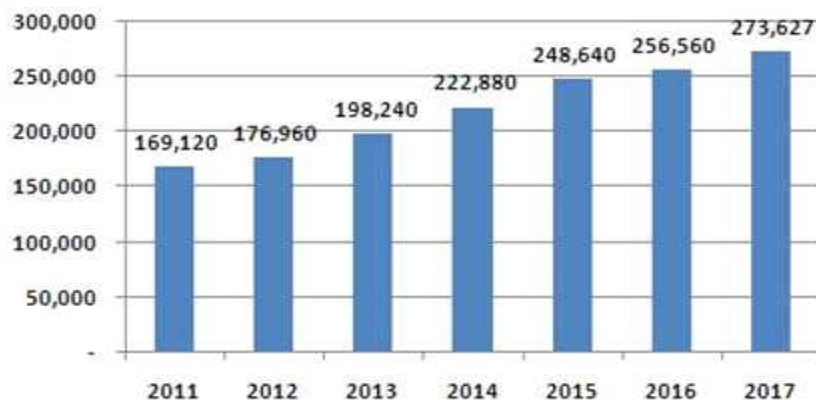
- 스트림 데이터를 분석하는 시장에서 경쟁하는 대표적 업체들로 EMC, 테라데이터, IBM, 오라클 등이 있음
- EMC의 경우, 대용량 데이터 저장이 가능한 스케일아웃 네트워크 스토리지(NAS) 아이실론을 포함해 하둡 배포판인 피보탈 HD, 하둡 기반 실시간 분산형 인메모리 통합 분석 플랫폼 피보탈 줌파이어 XD 등을 내세우고 있음
- IBM은 센서를 통해 적은 전력으로도 빠른 데이터 송출이 가능한 ‘MQTT(Message Queuing Telemetry Transport)’ 기술을 탑재하고 있는 메시지사이트라는 솔루션을 통해 실시간으로 대량 이벤트들을 처리하고 있음



<출처: EMC 빅데이터 분석 플랫폼>

2. 임베디드 SW 산업 분야에서의 빅데이터 기술

- 최근 산업 전 분야에서 IT 융합화가 빠르게 진행되고 제품의 지능화, 첨단화가 필수 요소로 되면서 임베디드 SW의 중요성이 급격히 증가되고 있음
 - 애플사는 iPhone, iPad 등 SW 혁신을 통해 글로벌 기업으로 성장한 반면, SW 경쟁에서 뒤쳐진 노키아, 소니는 글로벌 선도 기업에서 쇠락의 길을 걷고 있음
- 임베디드 SW는 자체로도 고속성장이 예상되고, 주력성장에도 기여함. 제조 산업의 전장화 비율이 증가하면서 임베디드 SW 산업의 부가가치 비중이 30%이상으로 확대되고 있음
 - '14년 임베디드 SW 국내시장은 약 22조원 모로 향후 5년간 매년 9.1%씩 급성장하여 '17년에는 27조원이 될 것으로 예상함
 - 분야별로는 자동차 6조원, 유무선 통신 4조원, 정보 가전기기 1조원, 조선 1조원 등으로 성장을 전망하고 있음
 - 산업통상자원부는 내년부터 2020년까지 6년간 자동차, 전자, 조선, 기계/로봇, 항공, 의료기기, 에너지 등 7대 핵심 분야에서 '100대 임베디드 SW 플랫폼'을 개발한다고 함



<출처: 국내 임베디드 SW 시장 규모 및 전망(New1)>

- 임베디드 SW는 많은 분야에서 각각의 특성에 맞게 적용되나 국내에서 주력하고 있는 산업분야의 개발현황은 다음과 같음

	과제	수행기관
자동차(2)	자율주행 자동차용 전방위 영상 모니터링 임베디드 SW 플랫폼	(주)와이즈오토모티브
	자율주행 자동차용 센서 활용 임베디드 SW 플랫폼	자동차부품연구원
전자(2)	스마트 가전 및 산업용 모션인식 임베디드 SW 플랫폼	포항공대 산학협력단
	원거리 기기 조작을 위한 RGB-D 카메라 기반 가상터치 인터페이스	(주)브이터치
조선(2)	NMEA2000 기반 선박 네트워크 임베디드 개발자 플랫폼 개발	전자부품연구원
	선박 센서 데이터의 입력차리 등에 활용되는 SW 플랫폼	뉴월드마리타임
기계/로봇(2)	3D 프린트의 실감 출력 제어용 임베디드 SW 플랫폼	(주)인레이저
	다양한 스마트 기기 연동이 가능한 로봇 미들웨어 개발	(주)로보케어
항공(1)	드론용 자율또는 원격 비행 임베디드 SW 플랫폼	(주)알티스트
의료기기(1)	의료 분야 웨어러블 디바이스용 임베디드 SW 플랫폼	하이버스(주)
에너지(1)	원격 전력 제어 및 충전용 임베디드 SW 플랫폼	(주)오픈비즈니스솔루션코리아

<출처: 2014년도 임베디드 SW플랫폼 개발과제 선정 결과(산업통상자원부)>

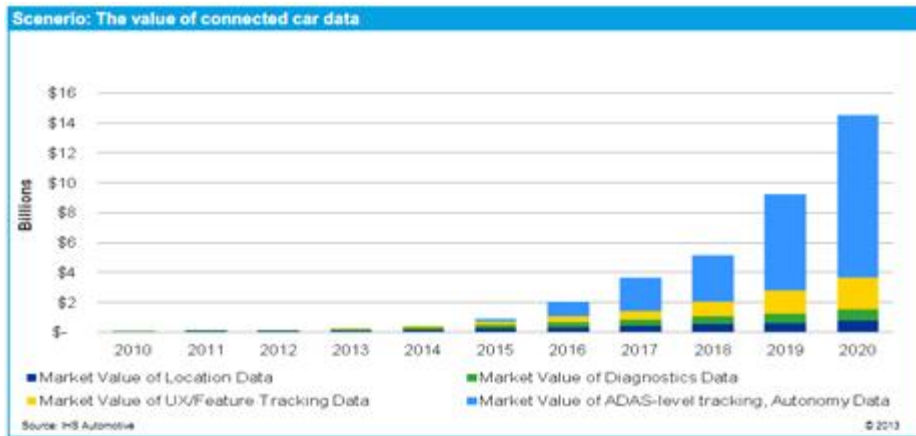
□ 제품의 고부가가치 및 첨단화를 유도하기 위해서는 임베디드 SW로부터 발생된 각종 데이터를 분석 처리하기 위한 기술이 필요함



<출처: 제조업 관련 빅데이터 기반 기술>

2.1 자동차 산업 분야

- 글로벌 분석 전문 업체인 IHS가 지난해 11월 자동차 시스템의 연결 기능과 높은 수준의 기술 통합으로 2020년까지 자동차 분야 데이터 자산이 145억에 달할 것이라고 전망된다는 내용의 보고서<Big Data in the Connected Car>를 발표함



<출처 : 커넥티드 카 데이터 자산(IHS Report: Big Data Will Represent Billions in Automotive)>

- 보고서는 따르면 2020년까지 전 세계 도로에 활성화된 연결 차량 센서는 1억 5,200만개에 이를 것이라고 전망함으로써, 이렇게 연결된 센서는 차량의 진단과 위치 추적, UX(User Experience) 데이터 수집, ADAS(Adaptive Driver Assistance System) 데이터 수집 등을 수행하는데, 약 2,600만개에 달하는 연결 차량으로부터 전송되는 데이터는 2013년에만 480TB, 향후 연결 차량 판매가 증가할 것이고, 엄청난 용량의 정보가 연결 차량으로부터 수집된다고 가정할 때, 2020년까지 이 분야에서 수집되는 데이터는 11.1PB 이를 것이라고 예측함

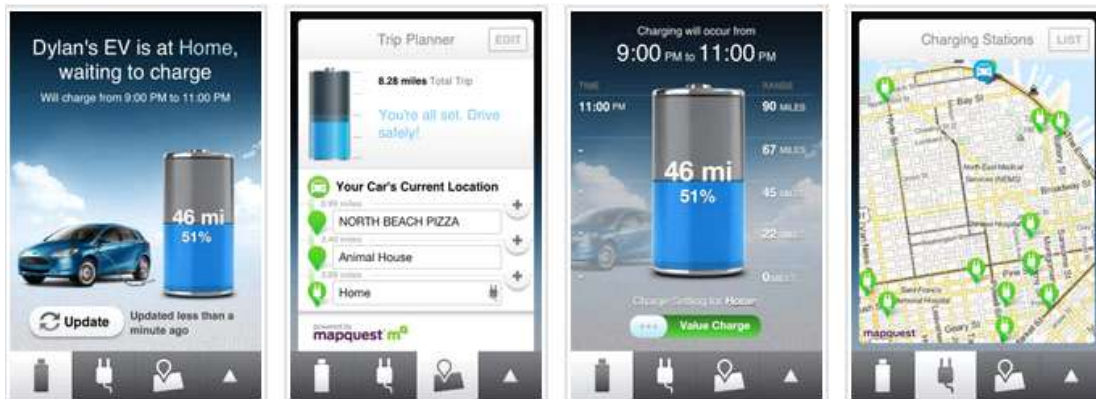
- 자동차 임베디드 SW는 엔진제어, 에어백 작동, 차간거리 자동제어 등 곳곳에 탑재되어 있음
- 자동차 산업은 자동차 부품의 전장화, 동력원의 전기화가 진행됨에 따라 세계 자동차 제조원가 중 전자부품 및 소프트웨어가 차지하는 비중은 2020년 35%, 2050년에는 50%까지 증가할 것으로 전망하고 있음
- 자동차산업의 핵심 부문은 기계부품 제작 및 조립 중심에서 IT제조, 소프트웨어, 첨단소재 중심으로 변화할 것으로 전망함



<출처: 미래형자동차기술(이데일리DB)>

- 에쿠스 자동차 한 대에는 '1270만 라인(줄)'으로 프로그래밍한 임베디드 SW가 탑재되어 있는데 이것은 A4 용지 400만장 분량임. 전자제어장치(ECU)와 같은 40여대 컴퓨터가 1000여개의 반도체 칩을 작동 및 제어하며 최적의 주행이 이뤄지도록 자동적으로 조절하는 기능을 포함하고 있음. 도요타의 자동차는 2015년 대당 1억 라인의 소프트웨어가 자동차에 들어갈 것으로 예상함

- 소프트웨어 동작 예로 자동차의 출발과 정지를 구현하는 것은 액셀러레이터를 뺏으면 엔진에 붙은 센서가 공기량과 공기의 온도를 감지하고, 엔진의 냉각 수온도 알아내고, 엔진의 회전속도와 배기가스 산소의 양도 감지해 냄. 이를 바탕으로 연료분사기(인젝터)에 얼마만큼의 연료를 분사하라는 명령이 내려짐으로써 최적의 연료 분사량으로 운전자가 원하는 속도를 내게 하는 핵심 기능이 임베디드 SW에 의해 동작됨
- 이렇게 자동차의 핵심 기능들은 소프트웨어들로 구현되고 있어 자동차 시장의 승부수는 각종 센서로부터 수집된 빅데이터를 오류 없이 실시간으로 분석해서 동작을 명령해야 함



<출처: Myford Mobile Screenshot
(<http://zeronova.kr/2013/08/19/ford-big-data/>)>

- 자동차 전자 제어 시스템은 외부정보를 받아들이기 위한 각종 센서 및 전자 장치를 부착하고 있음
- 포드자동차는 약 4백만 대 이상의 차량에 설치된 센서와 원격 어플리케이션 관리 소프트웨어를 통해 데이터를 수집함. 데이터 수집을 위해 포드가 차량에 설치한 센서는 차량 당 74개 이상. 여기에는 소나와 카메라, 레이더, 액셀러로미터, 온도 및 비를 감지할 수 있는 센서

등이 포함됨. 이렇게 설치된 센서는 엄청난 볼륨의 데이터를 생산. 플러그인 하이브리드 차량의 경우 매 시간 25 기가바이트의 데이터를 생산. 이는 차량의 공장으로부터 실시간 전송되고, 이렇게 분석된 데이터는 모바일 앱을 통해 운전자에게 다시 전달함

- Ford의 PHEV인 Ford Fusion Energi와 C-MAX Energi에서는 시간당 25 GB의 센서 데이터가 생성되고, 이를 이동통신망을 이용해 Ford의 클라우드 컴퓨팅 서비스에 모아서 데이터 처리를 한 후 다시 운전자의 스마트폰 앱(MyFord Mobile)으로 필요한 정보들을 전송함. Ford의 PHEV인 Ford Fusion Energi와 C-MAX Energi에서는 시간당 25 GB의 센서 데이터가 생성되고, 이를 이동통신망을 이용해 Ford의 클라우드 컴퓨팅 서비스에 모아서 데이터 처리를 한 후 다시 운전자의 스마트폰 앱(MyFord Mobile)으로 필요한 정보 제공함

□ 스마트카, 자율주행 자동차 등 차량 시스템의 지능화가 빠르게 진행되고 있음

- 스마트카¹⁾ 시대에 접어든 자동차 기술은 시스템 반도체 및 임베디드 SW 기술을 기반으로 전자제어 플랫폼이 핵심 기술로 부각되고 있음

1) 기계 중심의 자동차에 전기, 전자, 정보통신 기술의 융복합을 통하여 운전자의 부주의로 발생하는 교통사고의 획기적인 절감 및 탑승자의 만족(정보+편의제공)을 극대화한 차



<출처: 스마트카 시장 급팽창

(동아닷컴 <http://news.donga.com/3/all/20140507/63313838/1>)>

- 최근 구글, 르노, 메르세데스 벤츠, 아우디 등이 경쟁적으로 연구개발 중인 자율주행차(무인)차는 스마트카의 ‘꽃’이라 할 수 있음

사람처럼 판단하는 무인차 달릴수록 더 똑똑해진다

천장 레이저 센서가 사방 200m 보고 느끼면 도로상황 3D 지도로 변환해 주행 판단·제어
 습득 정보 축적... 상황에 맞는 대응법 학습 '물체'로 인식하다 '보행자 10명' 구분까지
 "거동 불편한 장애인·노약자들 이동에 도움" '컴퓨터가 모는 자동차'라는 불안감이 변수

구글 무인 자동차 어떻게 움직이나

숫자로 본 구글 무인차

- 탑승인원 — 2명
- 최고 속도 — 40km/h
- 주행 가능거리 — 160km
- 장애물 감지범위 — 200m
- 출시 예상시기 — 2017년

레이저 센서 360도 회전하며 사방 200m 이내 장애물 동시 감지해 3D 지도로 변환

카메라 신호등·보행자·자전거 등 장애물 및 교통 상황 감지

차량 전면부 보행자 보호 위해 휘어지는 신축성 있는 재질로 창문 및 전면부 제작

방향 센서 사람의 평형 감각을 담당하는 내이(內耳)처럼 차의 정확한 방향과 움직임을 감지

레이더 앞차와 거리 및 속도를 감지해 가속·감속을 결정

차 내부 출발·정지 버튼으로 차량 조작

운전석 주행 경로를 보여주는 화면과 짐 수납 공간

중앙 컴퓨터 차의 각 센서가 수집한 정보를 바탕으로 주행 제어, 조향 및 감속 등 주요 시스템은 이중 운영

위치 센서 휠 안쪽에서 바퀴 회전 감지해 차 위치 파악

배터리 전기 배터리 완전 충전 시 100km 주행 가능, 최고 시속 40km

구글 무인차가 본 도로 모습

구글 무인차가 본 도로 모습

- 보행자 (yellow cube)
- 자동차 (red cube)
- 자전거 (blue cube)
- 차량의 예상 주행 경로 (green line)
- 정지해야 할 위치 (red dashed line)
- 속도 줄여야 할 장애물 (green dashed line)

장애물 크기·속도·움직임 감지

정지 횡단, 신호등, 도로 구조 판단

움직이는 장애물의 속도, 방향 감지

보행자 발견시 차량 제어

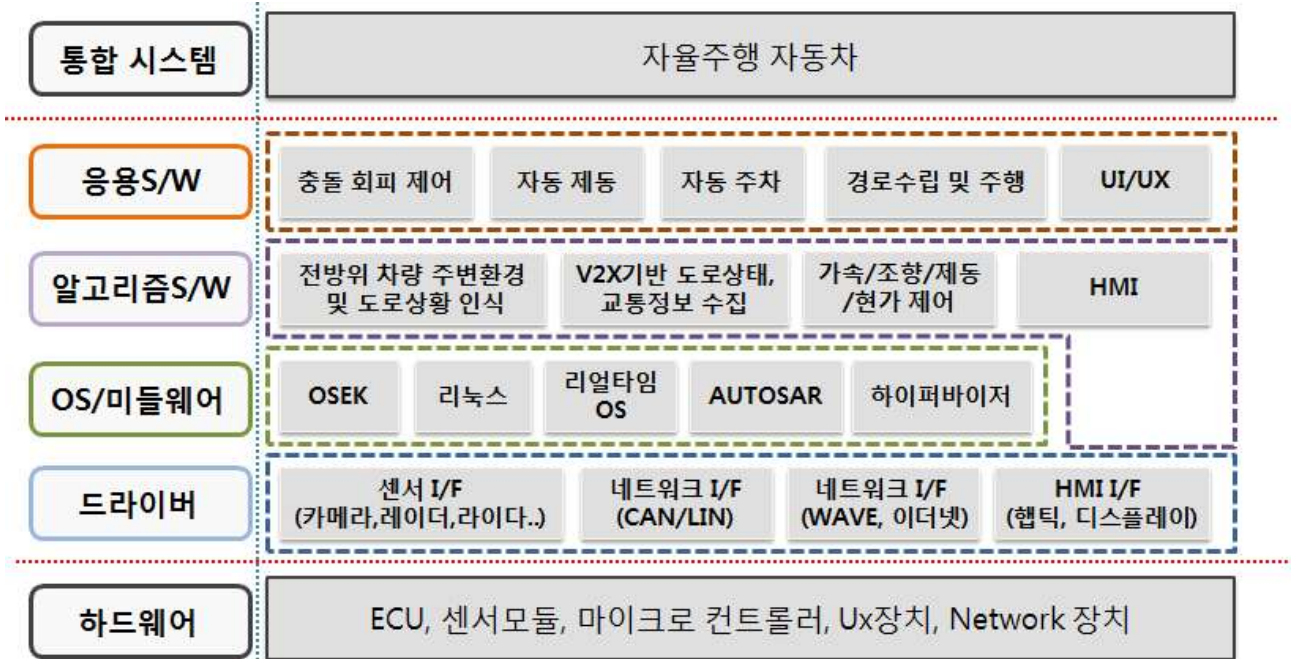
- 1 GPS(위성위치확인시스템)와 위치센서로 차량 현재 위치 파악
- 2 센서로 전방의 장애물 크기·속도·이동 경로 감지, 보행자 인식
- 3 보행자의 움직임·속도, 과거 학습내용을 바탕으로 멈추지 않고 건널 것이라고 예측
- 4 자연스럽게 차량 속도 줄여서 보행자에게 양보 의사 나타내며 정차

자료: 구글 그래픽 = 김현지 기자

<출처: 구글의 무인자동차 핵심 기술(출처: 조선닷컴)>

- 자율 주행 자동차는 스스로 주변 환경 정보를 수집, 인식하여 위험을 판단하고 회피하며, 주행경로를 계획하는 등 주행을 위한 운전자의 조작을 최소화 하고 스스로 안전주행이 가능한 자동차로 차량주변 전방위의 환경 및 도로상황을 인식하는 360도 고정밀 환경인식기술, 주

변 차량 및 도로의 상태와 교통정보를 수집, 전송하는 V2X 통신기술, 차량의 정밀한 위치를 파악하는 차량위치 측정 기술, 차량의 움직임을 결정하는 능동안전제어 기술 등으로 구성되어 있음



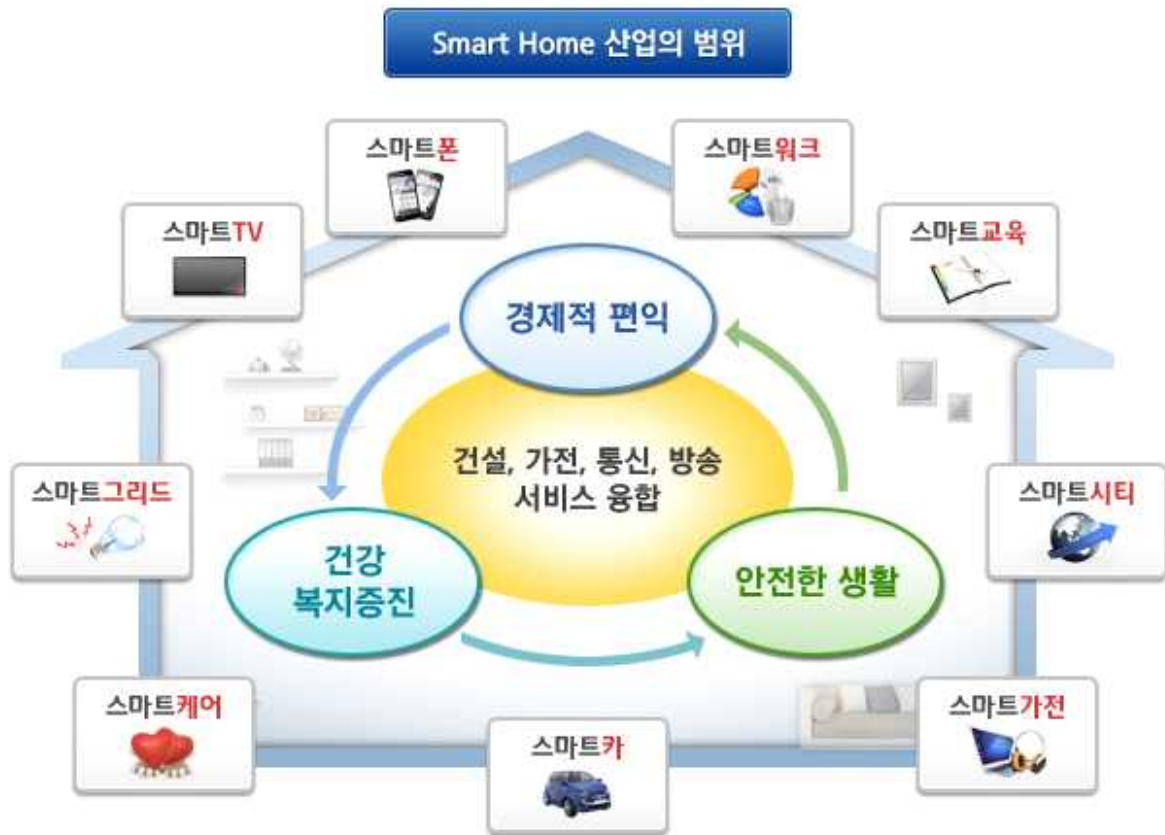
<출처: 자율 주행 자동차 임베디드SW 구조(임베디드 SW 경쟁력방안연구)>

- 구글이 2017년 상용화를 목표로 개발 중인 ‘무인자동차(self-driving car)’는 빅데이터, 센서, 인공지능, 자동제어 등 첨단 기술의 집약체. 실제 도로 주행을 통해 실시간으로 수집된 방대한 정보를 분석해 상황에 맞게 판단을 하게끔 분석함. 이러한 사례를 보면 자율 주행 자동차 핵심 기술은 자동차 제조 기술이 아니라, 자동차 외부의 여러 정보를 빠른 시간 안에 정확히 분석해 가장 알맞은 결정을 내리는 정보기술임. 구글은 빅데이터를 수집, 정리, 분석하는 핵심 기술을 이미 갖췄기 때문에 자동차 제조회사가 아닌데 무인 자동차 기술 개발을 선점하는 것이 가능했음

- 이러한 스마트카 산업이 발달할수록 발생하는 정보량은 많아지는데, 포드 자동차는 센서가 동작하면서 시간당 25GB의 데이터를 생산하는데, 센서의 개수 임베디드 SW 기술이 발달함에 따라 더 많은 정보가 생성될 것이고 이를 분석할 수 있는 빅데이터 기술이 중요해지고 있음
- 자동차와 관련된 빅데이터는 고객이 선택한 차량옵션, 현재까지의 수리 기록과 관련된 빅데이터, 주행 중 차량으로부터 수집된 주행 속도, 제동 장치, 에어백, 트랜스미션 등의 작동 기록 빅데이터, SNS 데이터, 고속도로 센서로부터 수집된 빅데이터 등 이러한 모든 각종 빅데이터들을 통신망 기술을 통해 융합하여 분석할 수 있는 것이 빅데이터 기술임
- 빅데이터 분석을 통해 차 소유주는 사전에 엔진 고장을 막음으로써 수리비를 절감할 수 있고, 자동차 업체는 부품 성능을 개선하는데 정보를 활용할 수 있음
- 실제로 볼보자동차는 빅데이터 기술을 이용해 기존에 50만대 판매 시점에 발견할 수 있었던 결함을 불과 1000대 판매 시점에 발견한 사례가 있음

2.2 전자 산업 분야

- 스마트홈(Smart Home)은 가전제품을 비롯한 집안의 모든 장치를 연결해 제어하는 기술임. 인텔리전트 빌딩의 각종 자동화기법(조명이나 온도 제어, 문과 창 제어, 보안시스템 등)은 가정 자동화에도 적용됨. 이 밖에도 안방극장 제어, 자동적이고 효율적인 에어컨, 방범시스템, 의료 시스템 접속 등이 포함됨

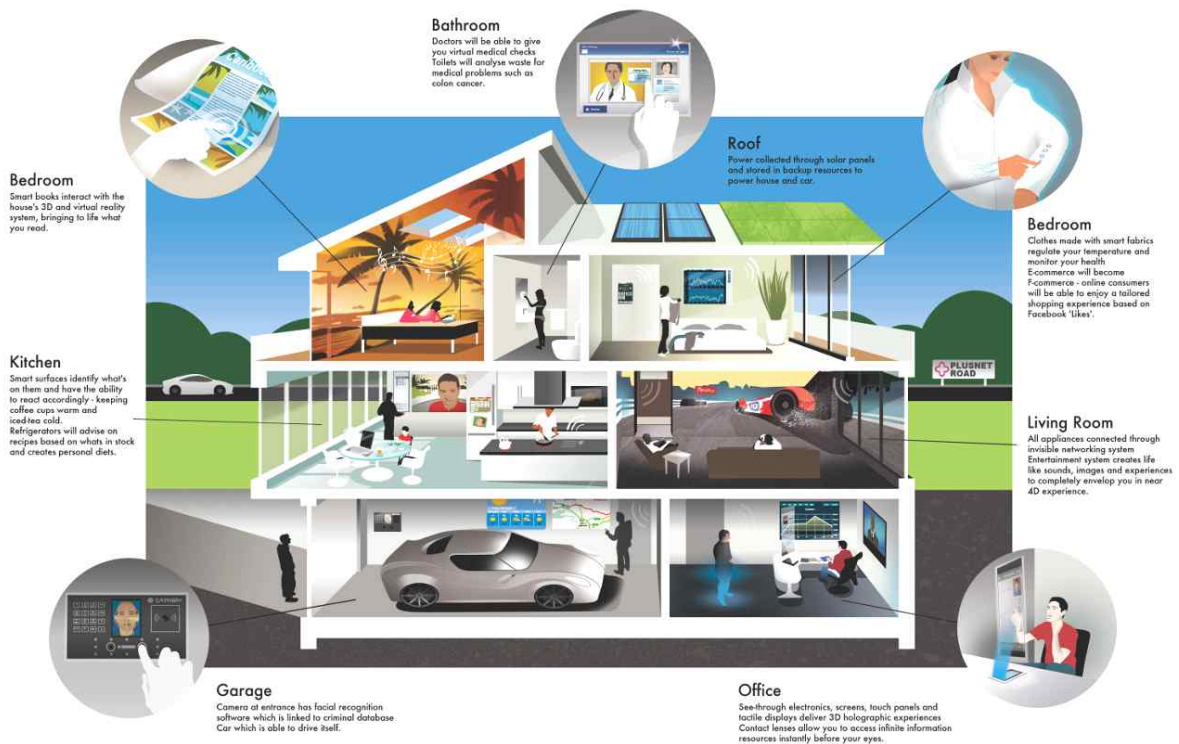


<출처: 한국스마트홈산업협회 (<http://www.kashi.or.kr>)>

- 스마트홈은 현재 추진되고 있는 유비쿼터스 환경과 결합하여 휴대용 기기로 언제 어디서나 가정과 관련된 일들을 처리하고 제어할 수 있게 되

는 꿈의 디지털 환경을 구현함. 정보가전 기기가 스마트홈의 단말 기능을 수행하기 위해서는 현재의 거의 모든 가전 기기에 관련 소프트웨어들이 탑재되어야 함. 임베디드 SW 관련 기업들이 정보가전과 스마트홈에 큰 관심을 가지게 되는 것도 바로 이런 이유 때문임

- 손 몇 번 움직이는 것만으로 내 주변의 모든 사물을 컨트롤할 수 있고, 어디서든 내가 원하는 정보를 얻을 수 있는 스마트홈을 만드는데 필요한 기술로 사물인터넷, 빅데이터, 클라우드가 핵심이 되고 있음



<출처: <http://community.plus.net/wp-content/uploads/2012/02/PlusnetHouseSmall.jpg>>

- 스마트홈은 단순히 기기의 연결을 통해 원거리에서 조작을 가능하게 하는 것을 포함하여 스마트홈 기기 스스로 데이터를 통해 상황을 분석하고 작동하여 이용자에게 알려주는 것이 가능해짐

- 예를 들어 외부 차고에 있는 차가 밤 동안 내린 눈의 양과 눈이 오는 날의 과거 교통상황을 고려하여 스마트폰의 알람을 조작하여 평소보다 30분 일찍 알람을 울리도록 할 수가 있음. 또한, 직장에서 퇴근 후 집이 가까워지면 아무런 명령 없이도 실내 조명등이 켜지고 로봇 청소기는 청소를 중단하며, 에어컨이 알아서 작동하는 등 집이 스스로 주인을 맞을 준비를 시작함

□ 스마트 가전은 인터넷 접속 기능을 내장한 똑똑한(smart) 냉장고·세탁기·오븐·청소기·에어컨 등의 가전제품을 통칭함. 새로운 조리법이나 세탁방법이 추가되면, 온라인에서 내려 받아 그 기능을 업그레이드할 수 있으며, 외부에서 원격으로 작동 상태를 조종할 수도 있음



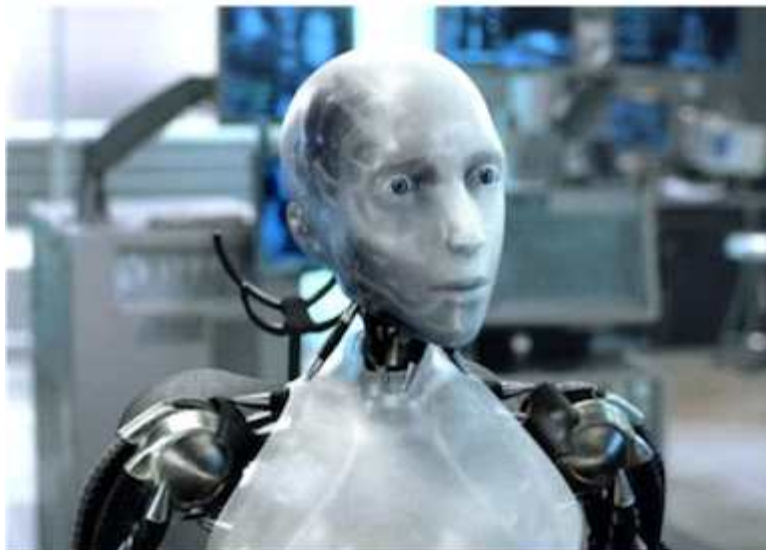
<출처: 삼성전자 스마트가전 개념도>

- 스마트 TV는 TV가 세탁기, 냉장고, 냉난방 시설, 출입 시설 등 가정 내 모든 기기와 통신하고 그 작동을 제어하게 되는 가정 내 미디어 허브뿐만 아니라 홈 허브로까지 그 기능의 확장을 추구할 것임
 - 스마트냉장고는 냉장고에 카메라와 LCD스크린을 장착한 후 인터넷을 통해 각종 식생활정보 정보 조회, 실시간 식재료 주문, 자가 절전, 보관 기한 관리, 냉장고 상태 확인이 가능함
 - 스마트에어컨은 실내외 기온과 날씨, 사람 수와 움직임 등을 내장 센서를 이용해 모니터링한 결과를 바탕으로 에어컨을 주위 환경에 맞게 자동 설정해주며, 스마트폰 내 위치정보를 이용해 집에 오기 전에 에어컨 전원을 켜거나 계속 쓰다보면 사용 패턴을 학습해서 자동 적용하는 기능도 갖추
- 스마트홈·스마트 가전의 센서를 통해 실시간 발생하는 데이터가 모여 빅데이터가 발생하게 되고, 이 데이터는 클라우드에 저장되어 실시간으로 수집, 관리, 분석되어 다시 기기로 전달되어 기기가 사용자에게 피드백을 주는 환경으로 구성됨

2.3 기계/로봇 산업 분야

- 지능형 로봇에서 임베디드 SW의 탑재는 가장 핵심 기술이며, 빅데이터 기반의 인지 컴퓨팅 기술 또한 지능형 로봇 플랫폼의 핵심임
- 로봇은 탑재된 임베디드 SW를 통해 주변 환경을 감지하고 데이터를 획득하고 데이터를 분석하여 자신의 행동을 결정함. 행동의 결과 데이터 또한 축적되어 이후 행동에 반영하는 과정을 반복하여 방대한 양의 데이터를 다루게 됨. 데이터가 많아질수록 인공지능의 능력이 늘어나서 빅데이터가 필요하게 됨

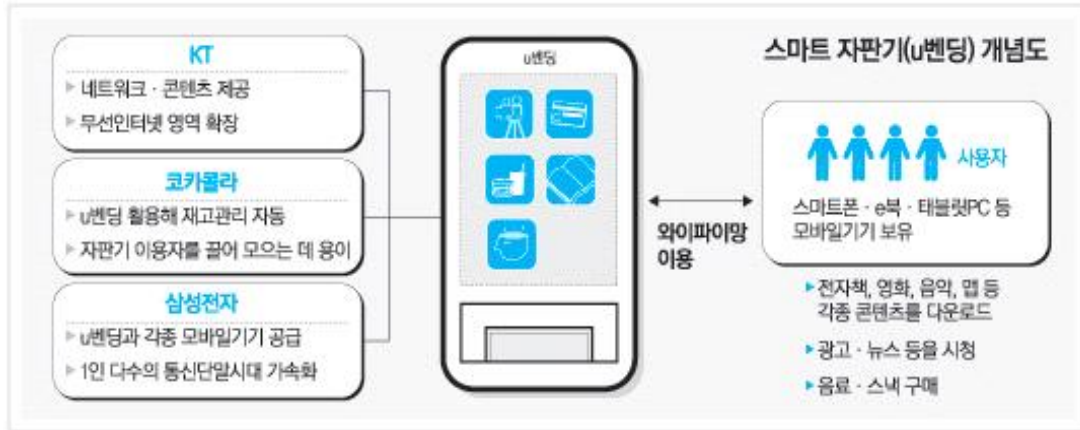
- 과거에는 로봇이 습득한 데이터를 분석하면서 학습을 했다면 빅데이터와 결합하여 이전까지 기록하지 못했던 수준의 데이터와 분석할 능력이 없었던 데이터들을 인터넷을 통해 모두 수집 및 분석할 수 있기 때문에 사실에 가까운 예측을 할 수 있음
- 대표적인 IT기업인 구글은 8개의 로봇회사를 인수함. 로봇은 스스로 탐지하고 학습하여 자신의 행동을 결정하고, 행동 결과를 학습하여 더 나은 행동을 만들어 낼 수 있음. 인공 지능을 만들어내기 위하여 구글 등 많은 기업들은 빅데이터를 활용함



<출처: 영화 '아이로봇'의 인공지능 로봇>

- 스마트자판기는 사물통신(M2M·Machine to Machine)과 임베디드 모바일 개념이 접목된 컨버전스 제품이자 양방향 능동형 기기임
- 코카콜라는 각 자판기들에 무선 인터넷 칩이 장착되어있어서 본사와 정보를 주고받아서 음료 소비량, 시간대별 판매량 통계, 인기 음료 등의 정보를 실시간 전송하며, 본사에서는 이 정보를 바탕으로 특정 지역에서 잘 팔리는 음료수를 판단하여 시간대별 음료 소비 패턴을 분석

하여 해당 지역에 적절한 홍보를 가능하게 함. 또한 재고 확인이 가능하여 소비자의 취향을 정확하게 파악하고 물류를 효율적으로 처리하는데 도움을 주고 있음



<출처: 매경신문 (<http://news.mk.co.kr/newsRead.php?year=2010&no=163935>)>

2.4 항공/의료기기 산업 분야

□ 미래 항공 시장을 주도할 무인항공기(Unmanned Air Vehicle)나 개인항공기(Personal Air Vehicle)를 위한 개발 기술은 IT융합 기술의 핵심으로 우리나라가 강한 경쟁력을 보이고 있는 분야임. 항공기내에 내장되는 항공전자 시스템인 임베디드 시스템과 이를 동작시키기 위한 항공기 임베디드 SW 기술 개발이 핵심 기술임

- 항공기 임베디드 시스템은 독립적인 프로세서들을 연결하여 동작하는 기존의 연합형(Federated) 모델을 벗어나 독립적인 모듈들을 통합하여 구성하는 통합 모듈형(IMA: Integrated Modular Architecture) 모델로 발전하고 있음

- 항공기 관련 데이터는 연간 25억 페타바이트가 발생하고 있으며, 항공 산업에서 빅데이터를 활용하면 항공기의 안전을 높이고 비용은 낮추는 효과를 얻을 수 있음
 - 비행기에 장착되는 제트 엔진 센서로부터 온도나 습도, 항공 압력 등의 데이터를 수집하여 다양한 기체 부품의 결함을 예측하고 적절한 시점에 예방적인 유지 보수 활동을 진행할 수 있음
 - 빅데이터를 통해 출발 및 도착이나 기상 상태, 기타 데이터 소스들을 활용하여 고객 관리 및 운항 지연 횟수, 비행 취소 건수를 줄일 수 있음
 - 항공기 엔진에 부착된 센서나 장비 작동기가 인터넷에 직접 연결됨으로써 수집 데이터를 즉시 전송 및 분석할 수 있어서 사물인터넷 기술을 이용하여 비행 계획과 운항시간, 연료 소비 등을 줄일 수 있음



<출처: HP, Wikibon>

- 아시아나항공은 운항통제센터(OCC: Operations Control Center)에서 빅데이터를 실무에 적용하고 있음. OCC는 항공기의 출발·도착 정보, 운항계획에 따른 비행기의 고도, 좌우편차, 연료 사용 여부 확인, 각국 공항의 활주로 상황, 기상 정보 등을 수집하여 1초 단위로 업데이트 및

분석을 통해 안전 운항을 위한 의사 결정 도구로 활용됨

- 의료기기는 IT, BT, NT 등의 다양한 기술과 의료 기술이 융합되어 언체, 어디서나 질병을 예방하고 진단하든 등의 의료 서비스를 제공받을 수 있도록 하는 기술로, 스마트 헬스케어에서는 데이터 수집, 데이터 전송, 데이터 관리/분석, 건강관리 예측 기술이 주요함
- 의료기기 산업 분야에서 임베디드 SW는 모든 기기에 장착된 센서뿐만 아니라 데이터 수집/처리 플랫폼, 네트워크 허브 기술, 관리 예측 시스템 등에 탑재가 되어 중요한 기술 중의 하나임
- 의료기기 산업에 있어서 빅데이터를 적용할 경우, 의료기관 별 진료방법, 임상기록, 수가기록, 관리기록, 진료의 효능 및 비용 데이터를 분석해서 보다 효과적인 진료방법을 파악하고 전국 의료 데이터를 연계하여 전염병의 발생을 예방할 수 있는 예측형 질병 관리 시스템과 개인별 의료 서비스 구축에 기여할 수 있음
 - 의료기기에 탑재된 임베디드 SW를 통해 환자에게서 발생된 수많은 센서 데이터들이 수집되고 수집된 데이터들은 인터넷에 있는 무한한 빅데이터들과 결합되어 빅데이터 분석 기법에 의해 환자의 상태를 파악하고 치료하며, 빅데이터 예측 기술을 이용하여 앞으로 생길 질병에 대한 예방까지 가능하게 됨

보건의료, 치료에서 예방·건강 관리로... 불치병까지 미리 막는 빅데이터

의료 빅데이터를 활용해 삶의 질을 높여주는 헬스케어 서비스들이 주목 받고 있다. 정보통신산업진흥원(NIPA)이 운영하는 IT서비스포털(www.itind.or.kr)에 게재된 국내외 보건의료 빅데이터 현황 및 과제 보고서에 따르면 치료에서 예방 중심으로 건강관리 확대에 따라 빅데이터를 접목한 보건서비스가 늘고 있다. 국민건강보험공단은 최근 건강보험 빅데이터와 소셜미디어 정보를 융합해 질병 발생을 예측, 위험을 미리 알려주는 '국민건강 주의 알림서비스'를 시작했다.

— 김달기 기자 kdelnews.com

보건의료 빅데이터 기술분야별 성장률

단위: %	전 세계	한국
데이터 수집단계	34.2	26.4
데이터 분류·시각화·연계	34.4	27
데이터 분석·예측·추론	39.9	27.8
데이터 시각화 표현	36.3	23
데이터 보안	32.8	20.8

보건의료 빅데이터 시장 성장률

단위: %	전 세계	한국
전체 보건의료	40.6	28.2
의료기관	37.2	25.1
보건기관	32.8	25.3
보건의료 기업	42	27.6

etview Plus
etview.etnews.com

맞춤형 의료서비스 활용 사례

▶ 국민건강 주의 예보서비스

공공기관, 국민건강정보, 민간 SNS 데이터

질병 위험 예측 시스템: 정형 데이터 연계/정형화 분석/마이닝, 정형 데이터 중량 분석, 비정형 데이터 중량 분석, 비정형 데이터 수집/마이닝, 비정형 데이터 분석/마이닝

국민건강 주의 알림 서비스: 주요 질병 모니터링 결과, 지역별/연령별 질병 위험 분석, 질병 관련 진료 동향 분석, 질병 관련 소셜 통합 분석

대국민 캠페인: 주요 질병 모니터링 결과, 예방 및 치료 정보, 지역별/연령별 질병 위험 분석 결과, 상시 건강 관리 정보

▶ 의약품 안전성 조기 경보 서비스

빅데이터 검색/수집, 지장, 부작용 정보 지장, 분석, 부작용 판단 임고리움, 조기 경보 시스템, 의약품 부작용 조기 경보 서비스

빅데이터 > 수집/지장 > 분석 > 조기 경보

미국 보행사 웰포인트의 IBM 슈퍼컴퓨터 '빛' 활용

- 웰포인트(Welpoint)는 빅데이터를 활용하여 복잡한 의료 절차 등에서 발생하는 비효율성을 제거
- 방대한 규모의 최신 임상정보, 환자-보험기록 등 의료 빅데이터를 빠르게 분석하여 가장 연관성이 높은 맞춤형 정보를 추출하고 신뢰도를 기반으로 의료진에게 제공, 정확한 의사결정을 지원함
- 빛은 60만건의 의학적 근거와 200만장에 달하는 전문서적, 2만5000여건의 사례 등을 분석

© 전자신문 인포그래픽·송광진 기자

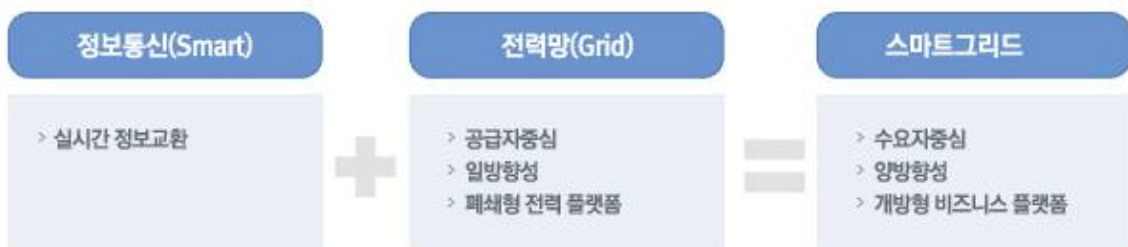
<출처: 전자신문 (<http://www.etnews.com/20140904000312>)>

2.5 에너지 산업 분야

- 스마트그리드는 전력망에 정보기술(IT)을 접목하여서 전력공급자와 소비자 양방향으로 실시간 정보를 교환하여 에너지 효율을 최적화하고 새로운 부가가치를 창출해내는 차세대 지능형 전력망임
 - 스마트그리드는 양방향 전력정보 교환을 하기 때문에 합리적인 에너지 소비를 유도할 수 있어서 고품질의 에너지를 제공받을 수 있음
 - 스마트그리드는 에너지의 생산부터 소비까지 전체 주기에 걸쳐 모니터링과 통제가 가능한 수준의 '연결'을 만들어내는 기술로써, 어떤 기기, 장치들이 어떻게 전력을 쓰고 있는지를 알 수 있으며, 에너지 관련 데이터를 일반가정에서 수집, 분석할 수 있는 환경이 갖추어짐
- 에너지 생산과 소비에 빅데이터가 접목되면 전력 공급자는 시장 수요에 대한 예측력이 높아져 언제 어떤 시기에 수요가 폭증할 수 있는지를 사전에 예측하여 그에 상응하는 조치를 사전에 취해 도시 전체가 정전되는

사태를 미연에 방지함

- 한전KDN은 스마트그리드 종합운영시스템에 빅데이터를 활용하여 제주 스마트그리드 실증사업 참여 및 수십 년간 축적된 전력ICT 전문 기술력을 바탕으로 빅데이터 기술을 전력계통 운영 및 스마트그리드 분야 등 다양한 신규 비즈니스 영역에 확대 추진 중임

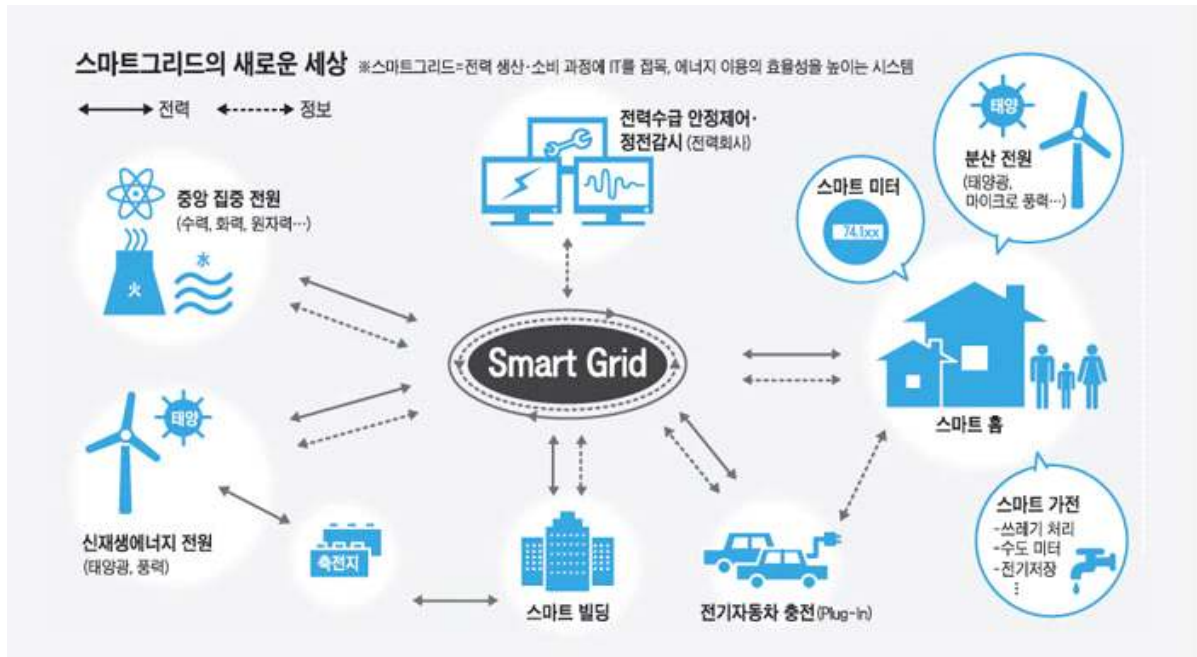


<출처: 제주스마트그리드실증단지 (<http://smartgrid.jeju.go.kr>)>

- 스마트그리드 개발은 새로운 IT(정보기술)와 OT(운영기술)의 결합뿐만 아니라 첨단계량인프라(AMI), 위상동조기, 지능형 가전기기, 마이크로그리드, 첨단 배전관리시스템, 원격 자산모니터링, 자가 치료(self healing)

네트워크 등이 말단 장치들에 의해 방대한 데이터 생산을 촉진할 것임. 데이터양의 폭발적 증가와 더불어 스마트그리드 이니셔티브는 데이터의 다양성 또한 확대될 것임. 예를 들어 일시적 형성 데이터, 공간구조 데이터, 거래 데이터, 스트리밍 데이터, 구조화 및 비구조화 데이터 등 다양한 종류의 데이터가 만들어질 것임

- 전력산업이 보안 및 신뢰성 우려 때문에 클라우드 컴퓨팅 도입에 있어 다른 산업 대비 후발주자의 모습을 보이겠지만, 클라우드 컴퓨팅의 솔루션들이 스마트미터, 빅데이터 분석, 수요 반응 조정 및 지리정보시스템(GIS) 등의 영역에서 성장하기 시작했다. 미국에서 유틸리티 클라우드 및 SaaS(software as a service)의 초기 도입이 투자자 소유 전력회사(IOU)들에 의해 이루어짐. 이 전력회사들은 단기 스마트그리드 시범사업을 추진하고 있으며, 생산시스템에 영향을 최소화하면서 시장 진입을 최대한 앞당기는데 관심을 보임
- 첨단계량인프라는 배전 자산 활용 결점 감지 능력을 향상시키고 에너지 시장에 소비자 참여를 촉진하기 위한 저지연(low-latency) 데이터의 통신 백본 역할을 제공함으로써 스마트그리드의 핵심 축 역할을 담당함. 다양한 시장 구조, 규제 체계, 편익에 대한 기대감 등은 첨단계량 인프라 기술 구조에 대한 다양한 소유 모델을 만들고 있음
- 유틸리티 자산의 분산된 특성은 보다 효율적인 자산 관리 및 노동인력 활용의 필요성과 결합하여 이동성과 통신 기술을 매우 중요한 투자 우선순위로 만들고 있음. 배전 네트워크를 보다 잘 관찰하고 발생하는 문제점에 대응하기 위해 추진하는 스마트그리드 사업에서는 낮은 대역폭의 SCADA(supervisory control and data acquisition)의 기능과 유사하면서 기능 및 장소별로 다양한 네트워크와 접근방식을 사용하는 사물대 사물(M2M, machine to machine) 기능을 필요로 함



<출처: 조선경제신문>

3. 임베디드 SW 산업과 빅데이터 기술 전망

- 임베디드 SW는 점점 더 디지털화와 멀티기능화 되고, 기기간의 네트워크 연결성이 확대되면서 앞으로 그 수요가 폭발적으로 늘어날 것이라는 점에서 시장 잠재성이 무궁무진함. 또한 실제로 임베디드 SW 산업에서도 일부 기업들은 이미 빅데이터 분석을 통한 제품 지능화를 통한 이윤 창출, 생산설비 개선을 통한 비용 절감 등의 효과를 이미 거두고 있음에 따라 임베디드 SW 산업과 빅데이터는 이미 밀접한 관계가 있다고 할 수 있음
- 웨어러블 기기와 사물인터넷, 무인 자동차 등 미래 산업에서 어마어마한 양의 생산되는 데이터 자체가 산업 분야의 큰 자산이 될 가능성이 큼. 이렇게 수집된 빅데이터의 분석을 통해 비즈니스 모델이 창출되고, 새로운 아이디어가 적용되는 제품 개발이 늘어날 것으로 예상됨. 이것은 지금보다 더 많은 수의 다양한 기기의 제품에 임베디드 SW가 탑재될 것을 예상함
- 이렇게 글로벌 기업들이 수행하는 스마트가전, 스마트자동차, 스마트팩토리 등 다양한 산업분야에서 발생한 빅데이터 수집 및 분석을 통해 ICT와 융·복합을 통한 새로운 가치를 창출하기 위해 다양한 신서비스, 신제품을 창출하기 위해 핵심적인 역할을 임베디드 SW가 수행하게 될 것으로 전망함
- 사물인터넷(Internet of Things)은 인터넷으로 서로 연결되어 모든 것들에 대한 정보 생성, 수집되고 공유, 활용되는 것으로 M2M(Machine to Machine), IoT(Internet of Things)를 거쳐 IoE(Internet of Everything)로까지 사용 범위가 확장되고 있음. 사물인터넷의 가장 기본은 IoT를 구동하는 기기들에 탑재된 임베디드 SW이므로 자동차산업, 유틸리티(전기

· 가스·수도), 의료분야, 가전분야 등의 통신 모듈 및 센서 기업 등에 사업기회가 창출될 것임

- 사물(기기)의 지능화로 인간의 삶이 보다 스마트하게 되는 사물인터넷 시대를 맞이하여 스마트 홈, 스마트 그리드, 헬스케어, 지능형 교통 서비스 등을 중심으로 서비스가 다각화될 것이며, 사물인터넷, 클라우드, 빅데이터 기술이 핵심이 될 것임. 사물의 데이터 수집 측면에서 핵심적인 역할을 담당하는 임베디드 SW는 데이터 처리·데이터 통신·데이터 보안 등의 핵심 기술을 개발하여 구글과 같은 IT기업이 독점하고 있는 사물인터넷 산업 시장에서 경쟁력을 가질 수 있는 방안을 모색해야 할 것임
- 특히 지금까지의 임베디드 SW 산업의 관심은 하드웨어 단말에 탑재된 소프트웨어나 기기들이 중심이었으나, 이들을 인터넷과 연결하는 사물인터넷 세상이 주류가 되면서 기기들 간의 통신이 가능한 M2M이 핵심이 되고 있음. 또한 기기들 간 주고받는 데이터들이 빅데이터로 수집·처리·분석·예측이 가능하므로 임베디드 SW 산업에서 빅데이터 기술 동향에 적극적인 관심을 보여야하며 임베디드 SW 산업에 적용 가능한 빅데이터 기술 확보에 주력해야 할 것으로 전망함

정한민 한국과학기술정보연구원 실장/책임연구원



□ 주요 활동 사항

- (前) 한국전자통신연구원 선임연구원
- (前) 다이퀘스트 연구소장
- UST 빅데이터전공 교수
- 한국콘텐츠학회 상임이사/편집위원/연구위원
- 한국정보과학회 컴퓨터지능소사이어티 이사
- 한국정보기술융합학회 부회장
- 융합정보학회 이사
- 한국인터넷정보학회 이사
- 한국빅데이터서비스학회 이사
- ISO/IEC SC32 전문위원
- 중앙공무원교육원 객원교수

저자

이미경 한국과학기술정보연구원 선임연구원

□ 주요 활동 사항

- (前) 한국전자통신연구원 연구원
- 한국과학기술정보연구원 컴퓨터지능연구실 선임연구원



저자

조민희 한국과학기술정보연구원 선임연구원

□ 주요 활동 사항

- 한국과학기술정보연구원 컴퓨터지능연구실 선임연구원



KESSIA ISSUE REPORT

발행처 :  **임베디드소프트웨어 · 시스템산업협회**
Korea Embedded Software and System Industry Association

서울시 마포구 월드컵북로 396 누리꿈스퀘어 비즈니스타워 8층

발행인 : 이 호 수

편집인 : 윤 동 섭

발행일 : 2014년 12월

전 화 : 02-2132-0755 팩스 : 02-782-1266

<http://www.kessia.kr>
