

# 빅데이터 활성화 정책 및 응용 사례

북경수, 유재수  
충북대학교

## 요약

다양한 정보 채널의 등장과 함께 빅데이터에 대한 중요성이 부각되고 있다. 본고에서는 빅데이터를 활용한 응용을 활성화하기 위해 국가별 빅데이터 정책을 분석하고 분야별로 빅데이터를 활용한 사례를 기술한다. 빅데이터를 활용하여 일상 생활에서 일반 사용자들이 사용할 수 있는 응용 서비스 원천 기술 및 서비스 기술을 개발하고 있는 충북대학교 빅데이터생활형서비스 연구센터(BLSRC)를 소개하고 센터에서 개발한 응용 서비스를 기술한다. 또한 빅데이터 사용을 활성화하고 실생활에 응용하기 위한 방안을 제언한다.

## I. 서론

다양한 채널 채널의 등장과 이로 인한 정보의 생산, 유통, 보유량의 증가로 디지털 데이터가 기하급수적으로 증가하고 있다. '07년부터 전 세계적으로 생성된 디지털 정보량이 사용 가능한 저장공간을 초과하기 시작하여 '11년 전 세계 데이터에 생성된 디지털 정보량이 1.8ZB에 달하는 제타바이트 시대로 진입하였다. 디지털 정보량의 기하급수적인 증가에 따라 대규모 데이터가 중대 이슈로 부각되며 빅데이터 (big data)라는 용어가 등장하였다[11][18][19].

스마트 단말 확산, SNS 활성화, 사물네트워크(M2M) 확산으로 데이터의 폭발이 더욱 가속화 되며 점차 빅데이터화가 가속화가 되고 있다. 새롭게 발생하는 데이터의 상당수는 비정형 데이터로서 기존의 기법으로는 분석에 한계가 존재가 존재한다. 빅데이터는 기존 데이터베이스 관리 도구의 데이터 수집, 저장, 관리, 분석하는 역량을 넘어서는 데이터 집합으로 거대한 데이터 양(volume), 빠른 데이터 유통 및 이용 속도(velocity), 데이터 다양성(variety)을 특징을 가진다[8][11][19][20][21]. 센서 데이터, 웹 로그 데이터, 소셜 데이터 등 새로운 데이터가 생성됨에 따라 정형 데이터에서 비정형 데이터에 활용이 증가되고

있다[2]. 빅데이터는 의미 있는 정보를 실시간으로 도출하여 트렌드 분석, 마케팅, 의사 결정 등 다양한 분야에 활용되고 있다. 센서 데이터, 웹 로그 데이터, 소셜 데이터 등 새로운 데이터가 생성됨에 따라 정형 데이터에서 비정형데이터로 분석 요구 변화하고 있다. 또한, 소비자 취향과 행동 등 다양한 변화 감지, 사람을 통하지 않는 신속한 의사 결정을 지원하여 산업 구조와 경쟁 양상을 변화시키고 있다[4].

글로벌 IT 기업은 빅데이터 시장을 대부분 선점하고 있으며 데이터 분석 중심으로 조직을 개편하고 역량 강화 및 기술 개발한다. IBM, Oracle 등 글로벌 IT 기업은 인프라, 데이터 수집·관리, 분석, 의사결정 지원 등 빅데이터 전 부문을 담당하고 있다[8, 18]. 글로벌 빅데이터 시장 규모는 '10년 32억 달러, '15년 169억 달러로 연평균 52% 고속 성장하고 있으며 '17년 324억 달러의 시장 규모를 형성할 것으로 예상하고 있다[12]. 국내 기업은 글로벌 기업에 비해 기술수준과 경쟁력이 취약한 수준이며 하둡(Hadoop) 등 빅데이터 분석 기술을 활용하는 서비스들이 증가하고 있으나 여전히 미흡한 상황이다. 국내 빅데이터 시장 규모는 '15년 2억 6,300만 달러, '17년 3억 3,600만 달러로 연평균 27.9% 성장할 것으로 예상된다[2].

미국, EU, 일본 등은 글로벌 경제위기, 데이터의 폭증, 데이터 활용 기술의 진화 등의 환경 변화로 인해 새로운 가치를 창출하고 신산업 육성을 위해 빅데이터를 집중 육성하고 있다. 해외 선진국의 경우 정부 차원의 빅데이터 활성화 전략, R&D 사업 육성, 법제도 개선 등의 정책을 마련하고 있다[15]. 국내의 경우 '11년 10월 빅데이터 활용을 통한 국가경쟁력 제고를 위해 빅데이터를 활용한 스마트 정부 구현방안을 마련하였으나 공공기관을 중심으로 한 시범 서비스가 위주로 진행되고 있다. 국내 정부 기관을 중심으로 공공데이터를 활용하여 범부처간, 정부·민간 융합지식을 도출하기 전략이 발표되었으나 공공 데이터 사용이 제한적이며 데이터 접근이 용이하지 않다. 국내 산업의 경우 주요 해외 기업들이 빅데이터 시장을 점령하고 있는 가운데, 해외 제품을 이용한 서비스 개발이 위주로 진행되고 있으며 국내 포털 사이트, 대기업체 등에서 빅데이터를 활용하기 위한 다양한 접근을 시도하고 있으나 일반 사용자가 생활에서 직

접적으로 사용 가능한 서비스 제공은 아직 미흡한 실정이다.

빅데이터는 현재의 상황 파악 및 가까운 미래를 예측할 수 있는 의미 있는 정보를 창출해야 하며 의미 있는 해석을 도출할 수 없다면 빅 데이터의 존재 의미가 없다. 현재 빅데이터는 금융, 행정, 문화, 사회, 예술, 보건 등 다양한 산업 분야에서 활용되고 있다. EU의 경우 15~20%의 공공 관리 비용 감소 효과와 2~4천 달러의 가치 창출을 할 것으로 예상하고 있다. 제조업 적용시 상품개발 및 조립 비용을 50%이상 절감하며 운용 자본도 7%이상 절감할 것으로 예상하고 있다[11][12][15].

본고에서는 빅데이터를 활성화하기 위한 국내외 정책 동향 및 빅데이터 활용 사례를 기술한다. 또한 충북대학교 빅데이터생활형서비스 연구센터(BLSRC)의 서비스 개발 사례에 대해 기술한다. II장에서는 영국, 미국, 싱가포르, 일본 등 빅데이터를 선도하고 있는 국가들의 빅데이터 활성화 정책을 기술하고 국내 빅데이터 정책을 기술한다. III장에서는 교통, 치안, 공공, 보건, 마케팅 분야에 빅데이터를 활용 사례를 기술한다. IV장에서는 빅데이터에 대한 생활형 서비스를 제공하기 위한 충북대학교 빅데이터 생활형 서비스 연구센터를 소개하고 센터에서 개발한 빅데이터 응용 서비스에 대해 기술한다.

## II. 빅데이터 활성화 정책

### 1. 영국

영국 경제경영연구센터에서는 '12년에서 '17년까지 빅데이터는 영국 내에서 약 2,160억 파운드의 경제적 이익을 창출할 것으로 예측하고 있다. 특히, 제조, 유통, 금융, 공공 부문 등 산업 전반에 걸쳐 효율성을 향상시키고 혁신을 창출시킬 뿐만 아니라, 빅데이터 기술 전문 신생 기업들의 증가에 따른 신규 일자리 창출까지 유발할 것으로 예상하고 있다. 영국 정부는 '13년 대량의 데이터를 활용해 일반 국민, 소비자, 기업, 교육 기관, 공공 부문 등 국가 전반의 새로운 가치를 창출하기 위한 목적으로 '빅데이터 역량 강화 전략'을 마련하고 빅데이터 전문 인력, 데이터 분석 및 저장에 가능한 인프라와 도구, 접근성과 활용성을 갖춘 풍부한 데이터 등 3가지 핵심 요소를 확보하고 있다[1][8].

영국 정부는 데이터 스토리지 시장의 글로벌 경쟁력 강화를 촉진하기 위한 정책을 전개하고 있으며 고성능 컴퓨팅, 클라우드 컴퓨팅 및 그 밖의 최신 기술을 통합해 대량의 데이터 수집, 저장, 분석, 시각화 등 빅데이터의 전 과정을 지원할 수 있는 e-인프라 시스템을 구축하고 활용을 확대하고 있다. 영국 정부는 데이터 공유를 통한 역량 강화의 차원에서 각종 연구 데이터

에 대한 활용 및 접근을 위한 지원 계획을 마련하고 있으며 '14년 초에 공공 과학 데이터 포럼을 개최해 연구위원회와 학계 간 연계를 확대하고 있다[3].

### 2. 싱가포르

싱가포르의 정보통신개발청(IDA)은 ICT R&D 정책 패키지인 정보통신기술 로드맵을 발표하여 각종 빅데이터 활성화 전략을 발표하였다[14]. IDA는 빅데이터 관련 전문 인력 양성, 실제 비즈니스 영역에 적용 가능한 빅데이터 분석 및 데이터 공유 서비스 개발, 빅데이터 솔루션 개발 사업자 지원 등 빅데이터 활성화를 위한 기반 조성 작업을 수행을 하고 있다.

경제개발청 주도로 정부와 기업의 경쟁력 강화를 위한 데이터 분석 연구소를 설립하고 국가안보조정국에서는 데이터 기반으로 위협 요소를 평가하는 RAHS시스템을 구축했다[3]. 또한, '12년 이를 운영하기 위한 기관을 설립했다. 이를 통해서 싱가포르는 해안 안전과 조류 인플루엔자 시뮬레이션을 실행하고 국가 위기 대응 능력을 향상시키고 있다[15].

싱가포르 정부 산하 국립연구재단은 싱가포르의 전략적 성장 분야로서 데이터 기반 비즈니스 분석 영역을 활성화시키기 위해 데이터 수집 및 분석 기술 개발 연구기관인 생활 분석 리서치 센터(LARC)를 운영하고 있다[14]. LARC는 SMU와 CMU 연구원들이 공동 참여하여 5개 분야에 대한 연구를 진행 중이다. LARC의 빅데이터 분석 및 실험 방법, 기술 플랫폼, 프로젝트 결과물 등을 싱가포르 경제개발청의 아시아 소비자 인사이트 연구소에서 적극 활용 중이다.

### 3. 일본

일본은 민간부에서 일부 빅데이터 활용 사례가 증가하고 있으나, 전반적인 활용 정도는 미국이나 유럽에 비해 낮은 상태이다[13]. 일본 정부의 데이터 처리와 관련된 기술 개발 및 정책은 2005년 이후 꾸준히 진행 중이다. 일본 정부는 정보통신 심의회에 빅데이터 활용에 관한 특별 그룹을 설치했으며, '12년 빅데이터 활용 정책의 기본방향을 정리한 빅데이터 활용 기본 전략을 발표했다[3]. 그 후 총무성은 '12년 차기 ICT 전략인 액티브 재팬의 5대 전략 중 한 가지로 빅데이터 이용과 활용에 의한 사회·경제 성장을 포함시켰다. 액티브 재팬 전략은 5개의 추진 전략으로 구성되어 있으며 궁극적으로 민간 분야까지 포함하여 빅 데이터의 이용 및 활용을 활성화하는 것으로 이를 위해 데이터 개발, 기반기술 연구개발, 표준화, 활용인재 확보, 사물 간 통신촉진, 규제 개선, 산학관 공동 추진, 성과 평가방법 등을 추진하고 있다[8].

총무성 정보통신심의위원회 ICT 기본 전략 위원회 산하 빅데이터 애드혹 그룹은 '12년 5월 빅데이터의 활용 기본방침을 발표하였다. 총무성의 빅데이터 정책은 정부와 지자체, 민간 사업자 등이 센서 및 M2M 통신을 통해 생성·수집되는 다종 다량의 데이터에 대해 사회 전반적으로 공유가 가능한 지식과 정보의 출현 촉진에 주안을 두고 있다. 이를 위해 개별 정보가 생성, 수집, 축적, 공개, 유통되는 전과정을 통합화함으로써 사회적 문제 해결과 경제 활성화 실현에 기여하고자 한다[14].

#### 4. 미국

미국 정부는 경제성장 및 일자리창출, 정부서비스 혁신 등 국가가 직면하고 있는 문제의 해결방안으로 빅 데이터의 가능성을 인지하고 빅 데이터의 새로운 가치정보와 데이터를 분석·추출·획득할 수 있는 역량확보가 시급하다고 판단하였다. '12년 3월에 대통령실 내 과학기술정책실은 국가 차원의 다양한 부처가 참여하는 2억 달러 규모의 빅 데이터 R&D 계획을 발표하였다[8]. 이 계획은 정부기관이 공공정보를 개방하고 빅데이터를 활용하여 공공서비스도 개혁하겠다는 의도를 담고 있다. 빅데이터를 활용하여 투명하고, 효율적이며 혁신적인 정부 서비스를 제공하겠다는 것이다[15].

미국은 빅데이터를 활용할 수 있는 기술개발 및 인력 양성하기 위한 구체적인 실천방안이 진행되고 있다. 세부 지원 프로그램들은 국방부(DoD), 국립보건원(NIH), 에너지부(DoE), 지질조사국(USGS), 국립과학재단(NSF) 등 총 6개 기관별로 84개 프로그램으로 구분되며, 총 2억 달러의 예산을 투입하였다. NSF는 지식 추론과 사회에 데이터 제공 및 관리를 위한 방법론 고안 등 장기적 전략을 구현하고 있으며 NIH는 세포, 분자, 영상, 건강과 질병 관련 데이터에 투자하고 있다. DoD는 대용량 데이터를 활용해 인지·지각·결정을 제공하는 자율 시스템을 구현하며 USGS는 지구 시스템 과학에 빅 데이터를 활용할 계획이다.

#### 5. 국내

대통령 소속 국가 정보화 전략 위원회는 '12년에 스마트 국가 구현을 위한 빅데이터 마스터플랜을 발표한 바 있으며 데이터의 창조적 활용을 통한 스마트 강국을 실현하기 위한 4개의 추진 목표를 제시하였다. 데이터의 비전과 목표를 달성하기 위하여 부처별 의견을 수렴하고 해외 사례를 기반으로 총 6개 분야의 16개 과제를 선정하였으며 그 중에 서 비교적 국민 혜택이 크다고 판단되는 과제를 중심으로 단계적으로 확산하고 있다[8]. '12년에는 빅 데이터 마스터플랜을 수립하였으며 '14년까지 우선 추진 활용 과제 및 적용 가능 대상 과제를 수행하고 '17

년까지 빅 데이터 활용을 확산하고 관련 기반 기술을 고도화할 계획이다. 빅데이터 마스터플랜안은 선제적 사회현안 해결, 맞춤형 대국민 서비스, 데이터기반 정책수립 지원, 빅데이터 기반 신성장동력 창출을 목표로 2017년까지 기반조성을 위해 4개 영역 12개 세부과제를 선정하고 정부와 민간이 총 5,000억 원의 투입을 계획하고 있다[13].

'13년 6월경 안전행정부는 빅 데이터로 미래 트렌드를 분석하고 국가의 미래전략을 과학적으로 수립해 나간다는 내용을 담은 '정부 3.0 추진 기본계획'을 발표하였다[8]. 각 부처에서 필요한 공공데이터를 연계·분류하고 공유·분석할 수 있는 기반을 마련하여 부처별 간에 시스템을 중복 구축하는 것을 방지하고 각 부처가 공동으로 활용할 수 있는 빅데이터 기반 시스템을 구축할 계획이다[9][10]. 국민권익위원회는 각 부처의 온라인 민원 정보를 수집, 관리, 분류, 분석 및 예측하는 민원 정보 분석 시스템을 '11년부터 운영하고 있으며 국립보건연구원은 중앙 인체 자원 은행을 운영하여 한국인의 질병 원인 규명, 진단·치료법 연구·개발에 필요한 인체자원을 확보 및 제공하고 있다.

### Ⅲ. 빅데이터 응용 사례

#### 1. 교통

서울시에서 야간 유동인구가 높은 지역들을 대상으로 운영된 심야버스는 귀가에 어려움을 겪고 있던 시민들의 문제를 효과적으로 해결한 정책이다[3]. 노선을 결정하고 시행하는 과정이 빅데이터 분석을 통해 이루어졌다. 2013년 3월 한 달간 시민들이 자정부터 새벽 5시까지 사용한 KT 휴대폰 전화 데이터 30억여 건의 빅데이터를 분석하였다. <그림 1>은 서울시 심야 유동 인구를 분석한 결과이다. 서울 전역을 1km 반경 1,250개의

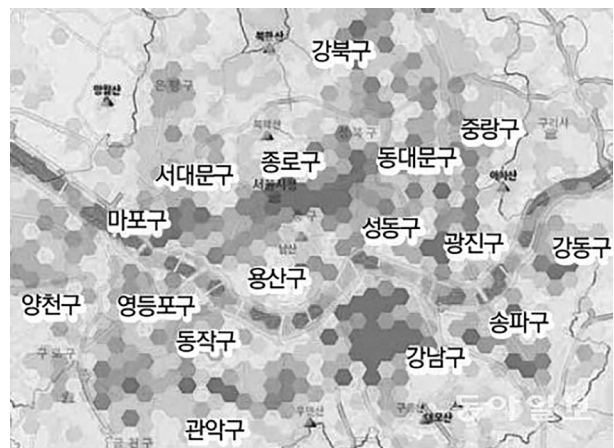


그림 1. 서울시 심야 유동 인구 분석





보하는 방식을 이용하고 있다. 또한, 과거 범죄에 대한 통계 정보를 제공하는 것과 달리 새로운 범죄 가능성 정보를 제공한다. 산타크루즈 경찰서는 2005년부터 범죄의 재발 가능성에 착안하여 기존 범죄자들의 SNS를 감시하고 동향을 파악하고 있다 [3][6]. 〈그림 5〉는 샌프란시스코의 범죄 발생내용을 나타낸 것이다. 6개월간의 테스트 결과, 예보 정확도가 71%에 달하였으며 범죄가 예보된 10곳 중 7곳에서 실제 사건이 발생하였다.



그림 5. 샌프란시스코의 범죄 지도

범죄자 유전자 데이터베이스를 통해 흉악 범죄자 검거를 위한 과학적 수사를 위한 기반이 마련되고 있다. 미국 FBI는 미제 사건 용의자 및 실종자에 대한 DNA 정보 1만 3,000건을 포함한 12만 명의 범죄자 DNA 정보를 저장하고 50개 모든 주와 연방 정부가 수집한 확정 판결을 받은 범죄자들과 일부 체포자들에게서 추출한 DNA 분석표로 구성하고 있다[6]. 내장된 DNA 분석 정보를 활용하여 '07년 45,400건의 범인 DNA 적중도를 달성하고 있으며 1시간 내에 범인 DNA 분석을 위한 주정부 데이터베이스 연계 및 빅데이터 실시간 분석 솔루션을 확보하고 있다.

미국 9.11 테러, 스페인 마드리드 열차 테러, 영국, 이탈리아 등 유럽 각국에서 발생한 테러 사건을 통해 국가 안보상에 대한 중요성을 인식하고 세계 국가의 출입국 관리 강화하게 되었다[17]. 싱가포르 출입국사무소는 출입국가, 수하물, 운송수단에 대한 통합적인 관리와 불법입국가 및 장기체류자 등에 대한 사전분석·예측을 실행하기 위한 솔루션 도입 필요 제기됨에 따라 위험을 사전에 파악하고 효과적인 관리를 진행하기 위해 WASP(Warehouse for Analysis, Statistics and Profiling)를 구축하였다. 싱가포르 국민 및 영주권자 등 출입국가 관련 데이터를 통합하여 불법 이민자 및 위험자들에 대한 데이터를 종합적으로 관리한다. 출입국 운영데스크에 위험 요소를 경보하고 위험자의 다음 경로 추적, 불법·위험행위의 사전 방지 및 불법 행위 발견 시 사법 처리 방안 등 추진하고 있다.

경기 침체 및 도덕적 해이로 인한 보험사기 증가되고 있다. KSTEC은 보험계약, 보험사고 등 보험 지급에 필요한 모든 정

보를 종합적으로 분석 보험사기 방지 시스템을 개발하였다[17]. 사기유형, 계약 및 사고유형을 개인, 보험모집인, 병원, 정비업소별로 구별하여 다양한 지표를 개발하고 혐의자 선정을 통해 가해자, 피해자, 동승자를 자동으로 추출하여, 사고 관련성, 공모 여부를 판단할 수 있는 연계분석시스템으로 체계적인 분석 가능하였다. 이를 통해 보험사기로 피해를 당하는 선의의 보험계약자 보호 및 우량 고객의 권익 보호에 기여할 수 있다.

### 3. 공공

글로벌 금융위기로 인한 재정 위기의 확산은 개인과 기업의 탈세에 따른 낭비성 재정 지출이 발생하고 있으며 미국의 탈세 금액은 '10년을 기준으로 저소득층 의료보장 총액을 초과하는 상황이 발생하고 있다. 이에 따라, 미국 국세청은 대용량의 데이터와 다양한 기술을 결합한 탈세 및 사기 범죄 예방 시스템 구축하였다[6][7]. 예측 모델링을 통해 과거의 행동 정보를 분석하여 사기 패턴과 유사한 행동 검출하는 사기 방지 솔루션을 제공한다. 또한, 계좌, 주소, 전화번호, 납세자 간의 연관 관계 분석 및 범죄자와 관련된 소셜 네트워크를 분석하여 범죄자 집단에 대한 감시 시스템 마련하였다. 통합형 탈세 및 정부사기 방지 시스템을 통해 연간 3,450억 달러의 세금 누락 및 불필요한 세금 환급 절감 효과를 거두었다[17]. 〈그림 6〉은 LA의 아동 복지 지원 서비스에 대한 사기 방지를 위해 분석할 결과이다.



그림 6. LA의 아동 복지 지원 서비스 사기 방지

최근 유로 재정위기의 확대, 세계 경기의 재침체 우려로 인해 국제 유가의 변동 폭이 매우 크게 변화하고 있다. 한국석유공사는 '11년에 데이터 분석 전문기업 SAS와 협력하여 유가 예보 시스템 개발하였다. 고유가에 따른 소비자 부담을 감소하기 위해 유가의 단기 미래 가격을 예측하여 제공하는 오픈넷 시스템 구축하고 국내 1,300여 개의 주유소로부터 수집된 휘발유 가격 정

보 제공한다[6]. <그림 7>은 오픈넷에서 제공하는 유가 가격 정보와 유가 예보 내용을 나타낸 것이다. 국제 유가를 기반으로 국내 정유사와 주유소의 판매가격을 추정하는 예측모델 개발하고 소비자들이 지역별, 상품별 및 5가지 변화 단계(상승, 소폭 상승, 포함, 소폭 하락, 하락)로 구분되는 시각화 자료를 제공한다.

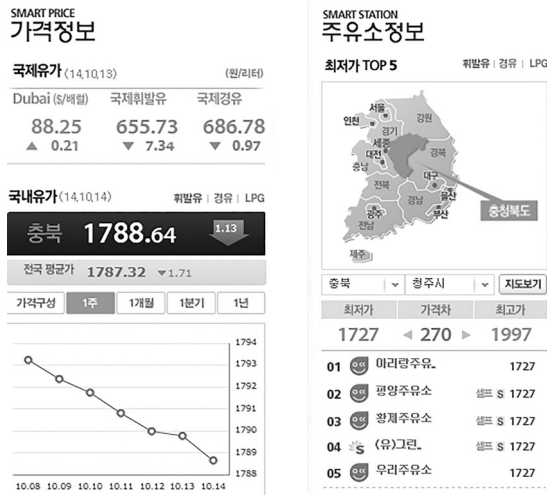


그림 7. 오픈넷의 가격정보 및 유가예보

대국민 서비스 활성화와 신뢰도 및 만족도 향상을 위해 범정부 차원에서 통합된 민원 서비스 연계 체계 구축 필요하게 되었다. 국민권익위원회는 민원의 체계적·과학적 분석이 가능한 민원 정보 분석 시스템을 이용하여 캘린더 형태로 제작하였다[17]. 과거 민원 발생 현황 등을 분석한 월별, 지역별 주요 민원 캘린더를 제작하여 사회적 이슈를 민원 지도 형태로 제공하며 월별, 지역별, 분야별 주요 민원 3~6개를 선정하여 캘린더 형태로 제작한다. <그림 8>은 민원 분석 시스템에서 민원 분석 결과를 지도를 나타낸 것이다. 배포한 월별 주요민원 캘린더와 지역별 주요민원 지도를 토대로 각 기관이 자체 민원유형을 추가하여 월별·지역별 주요민원 캘린더를 제작하고 이를 부서 게시판이나, 소속 직원의 책상 등에 부착하여 상시 활용할 수 있도록 함으로써 반복적으로 제기되고 있는 민원의 사전 예방과



그림 8. 민원 분석 시스템 및 민원 지도.

이에 대한 대책마련 등으로 행정력 낭비 방지와 신뢰행정이 구현될 것으로 기대하고 있다.

## 4. 보건

미국 국립보건원과 75개 기업 및 기관들이 파트너십을 통해 공동으로 진행한 1000 유전체 프로젝트의 일환으로 200TB의 유전자 정보 확보하고 있다. 국립보건원은 빅데이터 연구개발 이니셔티브 발표에 따라 1000 유전체 프로젝트를 아마존 웹서비스로 이전하여 연구원들은 유전자 데이터를 무료로 사용할 수 있으며, 사용한 만큼의 컴퓨팅 서비스에 대한 비용만 지불하고 서비스를 사용하고 있다[6]. 이를 통해 유전자 정보를 공유함으로써 새로운 질병에 대한 빠른 진단 서비스 제공 가능하며 난치병 및 불치병과 관련된 유전자 정보를 공유하고 분석함으로써 새로운 치료제 개발에 활용할 수 있다. <그림 9>는 아마존 웹 서비스를 통해 유전자 분석을 수행한 결과를 나타낸 것이다.

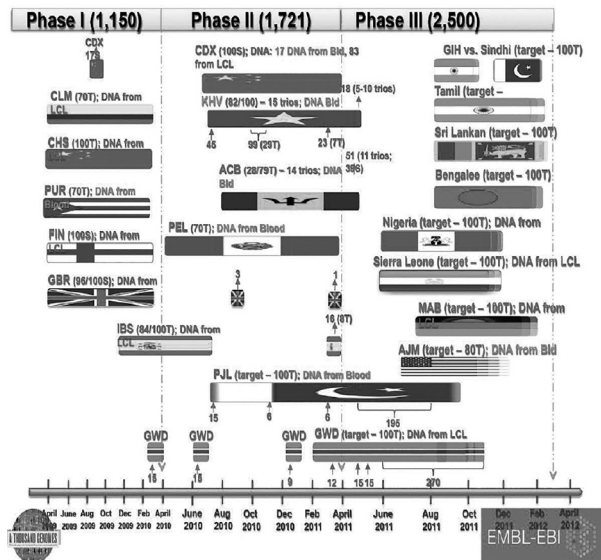


그림 9. 아마존 웹서비스의 유전자 분석

고령화 사회로 진입함에 따라 의료 비용의 급증함에 따라 병원의 정보화를 통해 효율화를 극대화하는 디지털병원에 대한 관심 증가하고 있다. IBM과 미국 의료보험사인 웰포인트(WellPoint)는 의사와 다른 의료진들이 진단과 환자 치료에 사용할 수 있는 애플리케이션 제공하고 있다[6]. 환자의 증상, 환자 면담 결과, 진단 연구 등 진료 내역에 대한 모든 정보를 저장하고 2억 페이지에 해당하는 자료를 검색 분석하여 3초 내에 결과를 제공할 수 있다. 이를 통해 빠르게 변화하는 진단 및 치료 방법을 서로 공유하여 환자의 진료 치료 만족도 증대시키고 불필요한 치료 및 진료를 줄여 환자 및 의료보험 회사의 불필요한



진료비 낭비를 방지한다.

미국 국립 의료원은 사용 중인 약에 대한 정보가 불분명할 때 약에 대한 정확한 정보를 확인하거나 알약에 새겨진 글자, 번호, 색깔, 모양, 크기 등 간단한 약에 대한 설명만으로도 정확한 약의 효능 및 정보를 제공하기 위해 필박스(Pillbox) 서비스를 제공하고 있다[6]. <그림 10>은 필박스에서 알약에 대한 정보를 검색한 결과를 나타낸 것이다. 필박스를 통해 알약의 제조사와 사용자간의 유기적인 쌍방향 상호작용을 통해 약에 대한 정보 제공하고 있다. 약에 대한 정보를 필요로 하는 사람의 대부분이 노년층이 많다는 점을 고려하여 누구나 쉽게 약을 검색할 수 있는 사용자 인터페이스 제공하고 있다.

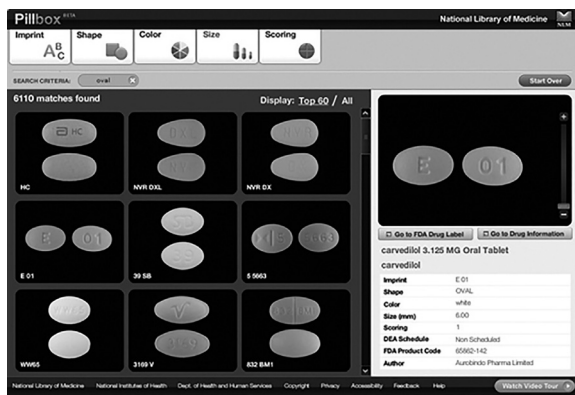


그림 10. 필박스 의약품 정보 서비스 검색

## 5. 마케팅

CRM을 이용한 고객 관리 시스템은 대부분의 대기업뿐만 아니라, 많은 중소기업에서 도입하여 실행하고 있다. 그러나 우수 고객 중심의 마케팅으로 인해 중상위층 고객들이 구매할만한 품목을 제대로 파악하지 못함으로써 직접적인 매출로 연계되지 못하는 상황 발생하고 있다[7]. 영국 테스코는 테스코의 클럽 카드를 이용한 고객의 구매행동 분석하여 다양한 고객의 특성에 맞춘 마케팅을 제공하고 있다. 고객 관리 프로그램을 통해 고객의 구매 행동을 분석하고 다양한 맞춤형 마케팅을 펼치면서 매출이 급속도로 증가하였다. 고객의 식품 구매 성향에 따라 20개의 특성군으로 분류하였으며, 지금도 매주 1,500만 건이 넘는 거래 데이터 분석하였다.

카탈리나 마케팅은 미국의 소비재(CPG) 기업과 유통업체를 고객사로 하는 글로벌 마케팅 기업으로 미국 내 2만 3,000개 이상의 매장과 1만 4,000개 소매 약국 그리고 전 세계 7,000개 매장에서 매주 2억 5,000만 건 이상의 트랜잭션으로부터 데이터 수집을 수집한다. 구매자의 구매 행동을 분석하고 예측하여

맞춤형 판매 시점 컬러 쿠폰, 광고 및 전국 소매 매장과 약국에 대한 정보지 제작하고 있다[7]. 또한, 고객이 계산대에서 계산할 때 고객의 구매 패턴을 다른 수천만 명의 구매 패턴과 비교 분석하여 가장 관심이 높을 만한 쿠폰을 즉석에서 발행한다.

미국에서 가장 오래된 동물원인 신시내티 동물원은 매출을 증가시키기 위해 수많은 마케팅 활동을 추진했으나 점점 매출 감소는 물론 정부보조금 축소에 따른 운영난에 직면하게 되었다. 이러한 문제를 해결하기 위해 데이터 분석 시스템을 구축하여 6개월간 입장객 행동 분석하여 새로운 마케팅 방법 도출하고 고객 만족도 증가시키고 있다[7]. 6개월 동안 동물원을 방문한 13,000명 이상의 관람객이 입장료 이외에는 돈을 쓰지 않는다는 사실 발견하고 이전에 시작한 모든 마케팅 활동을 중단하고 식음료와 상품 판매를 늘리는데 집중하였다. 또한, 아이스크림 가게 매출이 가장 더운 한낮이 아니라 해질 무렵에 가장 많이 판매 실적이 올라간다는 사실을 발견하여 운영시간을 2시간 연장하는 식으로 판매 방법 변화로 하루 2,000달러의 수익 창출하였다.

사용자와의 쌍방향 커뮤니케이션을 통해서 축적된 사용자에 대한 정보를 이용하여 사용자의 관심이나 선호도에 따라 오프라인 광고보다 사용자의 관심에 맞추어 광고 제공하고 있다. 야후는 온라인 행동 타겟팅 솔루션을 통해 사용자의 활동 및 관심도 패턴 분석하여 관심도가 높은 대상에게 관련 광고를 노출하는 광고 기법 개발을 개발하였다[7]. 검색, 광고 반응도, 콘텐츠 소비 방식 등 온라인상에서 이루어지는 이용자의 활동 및 관심도 패턴을 종합적으로 분석하고 특정 제품이나 서비스 카테고리에 관심이 높은 대상에게 해당되는 광고를 제공한다. 야후 일본은 2006년부터 모바일 이용자들의 온라인 행태에 관한 기록을 바탕으로 관심사를 분석하고, 각각의 모바일 이용자들에게 적합한 광고를 전송하는 모바일 기반의 맞춤형 광고 상용화하였다.

다양한 음식 관련 프로그램과 온라인 블로그, 레스토랑을 평가하는 스마트폰 어플리케이션의 증가에 따라 맛집 추천하는 다양한 방법을 제시되고 있다. 현대카드 가입자의 3개월 치 이용 실적 데이터를 분석하여 맛집 추천 서비스를 제공하는 마이메뉴앱 개발하였다[7]. 카드를 사용한 고객들의 이용 실적 및 정보를 이용하여 이용자에 맞춘 외식 추천 서비스 제공한다. 고객의 성별, 연령대, 직업, 재방문률 등의 항목을 세분화하여 파악한 정보에 리뷰를 결합하여 이용자에게 맞는 맛집을 추천하는 서비스를 제공한다.

## IV. 충북대학교 BLSRC의 서비스 개발

### 1. 소개

다양한 정보 채널의 등장과 이로 인한 정보의 생산, 유통, 보유량의 증가로 데이터가 기하급수적으로 증가함에 따라 빅데이터에 관심이 집중되고 있습니다. 빅데이터에서 의미 있는 정보를 실시간으로 도출하여 트렌드 분석, 마케팅, 의사결정 등 다양한 분야에서 활용되고 있으며 빅데이터 분석을 통해 현재 상황 파악 및 가까운 미래를 예측할 수 있는 의미 있는 정보를 창출할 수 있다. 충북대 ‘빅데이터 생활형 서비스 연구센터’는 스마트 공간에서 발생한 이종의 빅데이터를 처리하고 분석하여 개인에 직접 활용 가능한 생활형 서비스를 개발한다. 또한, 스마트 공간을 기반으로 빅데이터 처리 및 분석을 위한 IT 전문 인력 뿐만 아니라 학제간 융합을 통한 데이터 관리자 및 분석가

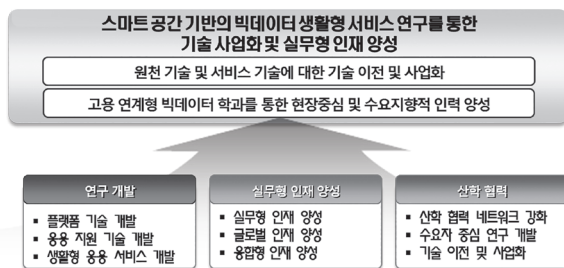


그림 11. 빅데이터 생활형 서비스 연구센터 목표

를 양성한다. <그림 11>은 빅데이터 생활형 서비스 연구센터의 목표를 나타낸 것이다

스마트 공간을 위한 스마트 라이프 케어 서비스를 제공하기 위해 스마트 공간에서 발생한 대용량의 빅데이터를 수집, 분산 저장 관리 및 분석을 위한 기반 플랫폼 기술 개발, 스마트 공간 서비스 구현을 위한 빅데이터 분석 및 시각화 기술을 개발하고 컴포넌트화, 스마트 공간에서 발생한 이종 빅데이터 처리 및 분석을 통한 스마트 라이프 케어 서비스를 개발한다. 또한, 본 센터에서는 스마트 공간에서 수집한 이종의 빅데이터를 활용하여 참여 기업체와 공동으로 빅데이터 기반의 스마트 라이프 케어 서비스를 개발하고 검증한다. 스마트 공간에서 수집한 다양한 유형의 데이터를 융합하고 분석하여 삶의 질 향상(smart life), 안전한 삶(safe life), 맞춤형 마케팅(tailored marketing)을 제공할 수 있는 스마트 라이프 케어 서비스를 제공한다.

스마트 라이프 케어 서비스는 스마트 라이프 서비스, 세이프 라이프 서비스, 맞춤형 마케팅 서비스를 제공한다. 스마트 라이프(smart life)는 스마트 장치, 사물 통신을 통해 수집한 정보를 통합하고 분석하여 사용자에게 보다 향상된 편의성을 제공하기

위한 서비스이다. <그림 12>는 스마트 라이프 케어 서비스의 제공 기능을 나타낸 것이다. 세이프 라이프(safe life)는 CCTV, 라이프로그 등의 정보를 이용하여 안전하고 건전한 일상생활을 수행할 수 있도록 서비스이다. 맞춤형 마케팅(tailored marketing)은 사용자 선호도 및 특성에 적합한 정보 제공을 위



그림 12. 스마트 라이프 케어 서비스

해 기업 비즈니스 환경 구축 및 차별화된 제품을 제공하기 위한 서비스이다.

### 2. 응용 서비스 개발 사례

기존 택시업계의 경력과 서비스의 질에 따른 소득 증가가 보장되지 않는 비구조적인 시장 구조로 인해 서비스 경쟁의지가 약화되고 있다. 고질적인 택시 기사의 불친절로 서비스의 질 하락에 따른 택시 이용 승객 불만 고조와 배회 택시 문화에 따른 에너지 낭비와 환경 오염 등 사회/경제적 악영향을 해결하기 위해 단골 택시 서비스를 제공한다. 단골 택시는 승객이 스마트폰 앱으로 주변의 콜택시를 직접 호출하는 무관제센터 방식 콜택시 서비스로 승객이 택시에 점수를 부여하고 이를 활용하여 택시를 호출하는 택시 서비스 평가를 제공한다. <그림 13>은 단골 택시 서비스의 실행 화면을 나타낸 것이다. 실시간 주변 택시 운행 상황 및 차량/기사 정보 파악 가능하며 차량정보, 탑승 후기를 통한 마음에 드는 택시 호출 가능하다. 승객과 기사 간 직거래 방식으로 콜센터 구축 등의 예산 절감할 수 있으며 승객



그림 13. 단골 택시



의 택시 평가 체계 구축으로 택시 문화의 선진화 실현 및 서비스 개선을 제공한다.

최근 다양한 범죄 증가로 인해 직장인(특히, 여성)과 아동이 피해를 받는 가능성이 증가되고 있다. 안심귀가 서비스는 실시간 데이터와 장소 데이터를 사용자에게 위험 정보를 실시간에 제공하며 다양한 위험 요소로부터 회피하여 안전하기 귀가할 수 있도록 도와주는 서비스이다. <그림 14>는 안심 귀가 서비스를 수행한 화면을 나타낸 것이다. 사용자의 시간대 이동 유형 및 행동 패턴을 분석하여 사용자의 상황에 따라 사고 발생 여부를 주변 지인에게 전달한다. CCTV, 뉴스데이터(사고관련, 재난, 범죄)를 수집하여, 시간대별 유동인구 분석, 실시간 데이터를 분석하여 사용자에게 제공한다. SNS데이터를 수집, 분석하여 지역별 위험 지역을 제공한다. 또한, 실시간 재난, 사고 데이



그림 14. 안심 귀가 서비스

터를 수집시 사용자에게 전송하며, 주변의 가까운 대피소, 관공서 위치 정보를 전달한다.

보건복지부 통계 자료에 따르면 국내 65세 이상 노인 중 9%에 해당하는 52만명이 치매환자로 추정되고 있다. 치매노인 모니터링 서비스는 고령화 시대에 병원 및 가정에서 활용 가능한 치매노인 모니터링 기능을 제공한다. 과학적인 일상생활 패턴 분석을 통한 개인 맞춤형 케어 서비스 제공하며 침실 및 화장실 낙상 분석을 통한 위험 분석 서비스 제공한다. <그림 15>는 치



그림 15. 치매노인 상태 분석

매 노인의 행동 패턴을 분석할 결과이다. 치매노인의 일상생활 주기 및 활동 내역 분석하고 과거와 현재의 생활패턴을 비교 분석하여 환자의 이상 패턴을 감지한다. 병원 내부 센서데이터를 수집 및 분석하여 침실 및 화장실의 낙상 판단하고 환자의 낙상 패턴 분석을 통한 낙상 위험 상황에 조기 대응할 수 있다. 다중 센터 데이터를 분석하여 안전지역 이탈 및 제한구역 출입시 실시간 알림 및 녹화 기능을 제공하여 치매 노인의 안전한 생활을 영위할 수 있도록 한다.

다양한 교통 수단의 발달과 함께 기상, 교통 통제, 교통 혼잡 등으로 인해 교통 정체가 발생하고 있다. 이와 함께 대중 교통을 이용자는 시내·시외버스, 지하철 등 다양한 대중 교통 정보들을 연계하여 하나의 통합된 형태의 서비스로 제공받고자 한다. 최적 경로 안내 서비스는 실시간 교통정보와 사용자의 현재 상황을 수집 및 연관 분석하여 상황에 맞는 최적화된 경로를 사용자에게 제공한다. 차량 운행 정보와 실시간 상황 정보를 고려한 경로 안내 서비스를 제공하고 실시간 교통 정보를 활용하여 대중 교통 환승 구역 안내한다. 이를 위해 차량 운행 정보, 실시간 교통 상황, 교통 데이터 분석하고 장소, 유형, 시간대별 분석한다. <그림 16>은 대중 교통에 대한 최적 경로 환승 구역을 나타낸 것이다. 날씨, 시간 등 교통에 영향을 주는 요소들을 분석하여 교통 상황 예측하고 과거 분석 결과 및 현재 교통 흐름이 결합된 교통 상황을 예측한다. 분석된 교통 정보는 공공기관에



그림 16. 최적 경로 안내 서비스

서 제공하며 공공기관의 정책 반영 및 유사 대민 서비스에 연계 활용할 수 있다.

## V. 결론

본고에서는 빅데이터에 활용을 증대시키기 위한 국가별 정책 분석 결과를 제시하고 유형별 빅데이터 활용 사례를 제시하였다. 또한, 충북대학교 빅데이터생활형서비스 연구센터에서 스

마트 라이프 케어 서비스에 대한 개발 사례를 기술하였다. 현재 국내 산업의 경우 주요 해외 제품들이 빅데이터 시장을 점령하고 있는 가운데 일부 포털, 대기업체 등에서 빅데이터를 활용하기 위한 다양한 접근을 시도하고 있다. 그러나 아직까지 일반 사용자가 생활에서 직접적으로 사용 가능한 서비스 개발 사례가 미흡한 실정이다. 빅데이터에 대한 사용을 활성화하기 위해서는 데이터 활용을 위한 국가적 지원책이 요구되며 서비스 개발을 위한 인프라 구축 및 데이터 연계를 위한 지속적인 협력 및 지원이 요구된다. 공공 기관에서 보유한 데이터를 활용하여 시범 서비스 개발을 진행하고 있으나 정부 기관 또는 지자체에서 보유한 데이터 활용 어려운 실정이다. 이를 해결하기 위해서는 민간 데이터와 공공 데이터의 연계·활용을 위한 체계 및 기술 확립이 필요하다. 또한, 공공데이터의 단계적 개방으로 민간의 가치창출 및 기업 활동 강화 지원이 요구된다. 데이터 활용 및 수집에서 정보 공개 및 개인 정보 활용을 위한 법적, 제도적 한계가 발생한다. 개인 정보 삭제 또는 프라이버시 침해를 하지 않는 상황에서 빅데이터를 활용하기 위한 제도 개선이 필요하며 관련 법령 개정이 필요하다.

## Acknowledgement

본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학IT 연구센터육성 지원사업/IT융합고급인력과정지원사업(NIPA-2014-H0301-14-1022), 교육부와 한국연구재단의 지역혁신 인력양성사업(No.2013H1B8A2032298), 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환으로 수행하였음(14-824-09-001, 실시간 대규모 영상 데이터 이해·예측을 위한 고성능 비주얼 디스커버리 플랫폼 개발). 또한 빅데이터 활성화를 위해 한국정보과학회지에도 동시에 게재함.

## 참 고 문 헌

- [1] “영국 정부의 빅데이터 역량 강화 전략”, 정보통신산업진흥원 동향 보고서, 2014.
- [2] 박세환, “빅데이터 기술 및 시장 동향”, 정보통신기술진흥센터 주간기술동향, pp.15-24, 2014.
- [3] 성지은, 박기량, “빅데이터를 활용한 정책 사례 분석과 시사점”, 과학기술정책연구원 과학기술정책, 제24권, 제2호, pp.94-106, 2014.
- [4] “빅데이터(Big Data) 활용단계에 따른 요소기술별 추진동향과 시사점”, 한국방송통신전파진흥원 방송통신기술 이슈&

- 전망, 제10호, pp.1-19, 2013
- [5] <http://opengov.seoul.go.kr/section/1520565>
- [6] “Big Data 글로벌 10대 선진 사례 - 빅데이터로 세상을 리드하다”, 한국정보화진흥원, 2012
- [7] “너 나은 미래를 위한 데이터 분석 Big Data 글로벌 선진 사례 II”, 한국정보화진흥원, 2013
- [8] 배상태, “과학기술정책 네트워크 분석 기반의 빅데이터 활용 방안 연구”, 한국과학기술기획평가원 정책기획 보고서, 2014.
- [9] “정부3.0 빅 데이터 활용 확대 방안”, 안전행정부, 2013
- [10] “정부3.0 세부추진계획”, 안전행정부, 2013
- [11] 이진형, “데이터 빅뱅, 빅 데이터(BIG DATA)의 동향”, Journal of Communications & Radio Spectrum, 통권 47호, pp.43-55, 2012
- [12] “빅데이터 산업의 현황과 전망”, KISTI Market Report, 2013
- [13] 배동민, 박현수, 오기환, “빅데이터 동향 및 정책 시사점”, 정보통신방송정책, 제25권, 제10호, pp.37-74, 2013
- [14] “싱가포르의 빅데이터(Big Data) 활성화 정책 분석”, 해외 ICT R&D 정책동향, 2013년 07호, 2013.
- [15] 최진명, “빅 데이터 시대(Big Data Era)의 데이터 활용과 전략”, 한국지역정보개발원 지역정보화 동향 분석, 제3호, pp.1-11, 2012
- [16] “일본 빅데이터 정책 추진 현황 분석과 국내 시사점 I”, 한국인터넷진흥원 글로벌 정보통신(ICT)·방송 Weekly Issue, 2014
- [17] “빅데이터로 진화하는 세상 Big Data 글로벌 선진 사례”, 한국정보화진흥원, 2012
- [18] M. Saecker and V. Markl, “Big Data Analytics on Modern Hardware Architectures: A Technology Survey”, Proc. European Business Intelligence Summer School, pp.125-149, 2012
- [19] J. S. Ward and A. Barker, “Undefined By Data: A Survey of Big Data Definitions”, The Computing Research Repository, 2013
- [20] M. Chen, S. Mao, and Y. Liu, “Big Data: A Survey”, Mobile Networks and Applications, vol.19, no.2, pp.171-209, 2014
- [21] “Big Data: Beyond the Hype: Why Big Data Matters to You”, BY DATASTAX CORPORATION White Paper, 2013

## 약 력



복 경 수

1998년 충북대학교 수학과(이학사)  
 2000년 충북대학교 정보통신공학과(공학석사)  
 2005년 충북대학교 정보통신공학과(공학박사)  
 2005년~2008년 한국과학기술원 전산학과 Postdoc  
 2008년~2011년 (주)가인정보기술 연구소  
 2011년~현재 충북대학교 정보통신공학과 초빙부교수  
 관심분야 : 데이터베이스 시스템, 위치기반서비스,  
 모바일 P2P 네트워크, 소셜 네트워크 서비스,  
 빅데이터 등



유 재 수

1989년 전북대학교 컴퓨터공학과(공학사)  
 1991년 KAIST 전산학과(공학석사)  
 1995년 KAIST 전산학과(공학박사)  
 1995년~1996년 목포대학교 전산통계학과 (전임강사)  
 1996년~현재 충북대학교 정보통신공학부 및  
 컴퓨터정보통신연구소 교수  
 2009년~2010년 캘리포니아주립대학교 방문교수  
 관심분야 : 데이터베이스시스템, 빅데이터,  
 센서네트워크 및 RFID, 소셜 네트워크 서비스, 분산  
 객체컴퓨팅, 바이오인포매틱스 등