

창조경제 실현을 위한 2013 빅데이터 국내 사례집

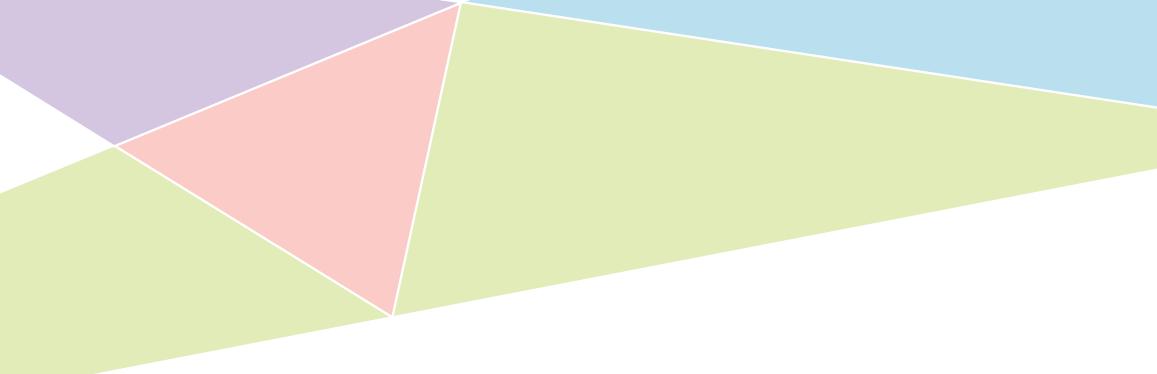


미래창조과학부



한국정보화진흥원





CONTENTS

I

빅데이터 시대

01 빅데이터란 무엇인가?	06
02 빅데이터 사례집의 목적 및 구성	11

II

빅데이터 활용 사례 Part 1. 미래창조과학부 시범사업 6종

01 빅데이터 분석을 통한 심야버스 노선 정책 지원	14
02 국민건강 주의 예보 시범 서비스 구축	18
03 의약품 안전성 조기 경보 서비스	22
04 심실부정맥 예측, 입원병상 최적화 등 보건의료 서비스	26
05 소상공인 창업지원을 위한 점포평가 서비스	30
06 스마트 뉴스 제공 서비스	35

III

빅데이터 활용 사례 Part 2. 공공사례 11종

01	빅데이터 기반 영주사과작황정보 서비스 (경상북도, 미래창조과학부)	42
02	기상 빅데이터 분석을 통한 효율적인 위험기상 예측 (기상청)	49
03	안전한 금융서비스를 위한 여신상시감시시스템 (금융감독원)	52
04	특수질병 및 질환 예방 및 관리 (보건복지부)	55
05	통합형 재난안전 정보 네트워크, 스마트 빅 보드 (안전행정부)	59
06	맞춤형 정보 유통 서비스 (안전행정부)	62
07	지리정보 통합한 지리적 프로파일링 시스템 구축 (경찰청)	65
08	국가공간정보 통합데이터베이스 구축 (국토교통부)	68
09	고용정보통합분석시스템 구축 (한국고용정보원)	71
10	역외탈세 정보공유 (관세청, 국세청)	73
11	빅데이터 분석을 통한 2013 국민 인식 변화 분석 (문화체육관광부)	76

IV

빅데이터 활용 사례 Part 3. 민간사례 13종

01	빅데이터 분석을 통한 베노플러스겔 마케팅 (유유체약)	82
02	SSD 온라인 버즈 마케팅 (삼성전자)	85
03	C.S.I 프로젝트 (신세계몰)	90
04	고객 맞춤형 서비스를 위한 빅데이터 센터 (신한카드)	93
05	소비자 맞춤형 수요 매칭 (FNC 코오롱)	96
06	고객 맞춤형 매장 배치 (현대백화점)	99
07	맞춤형 고객 상품 추천 서비스 (GS 삼)	102
08	맞춤형 여행 컨설팅 서비스 (롯데관광)	105
09	자동차 고객관계 관리 프로젝트 (현대자동차)	107
10	스마트 미디어 솔루션 서비스 (삼성전자)	110
11	모바일 빅데이터 맞춤형 서비스 (NHN)	114
12	스마트 통합물류 시스템 (CJ 대한통운)	117
13	언더라이팅의 예측모델 활용 방안 (푸르덴셜생명보험)	120

I

빅데이터 시대



01 빅데이터란 무엇인가?

02 빅데이터 사례집의 목적 및 구성

01

빅데이터란 무엇인가?

❶ **빅데이터란 데이터를 구성하고 있는 물리적 하드웨어로부터 시작해 이를 기반으로 하는 애플리케이션과 소프트웨어를 포함하는 거대 플랫폼**

- 단순한 데이터의 크기가 아니라 데이터의 형식과 처리 속도 등을 함께 아우르는 개념으로 기존 방법으로는 데이터의 수집, 저장, 검색, 분석 등이 어려운 데이터를 총칭해서 일컫는 용어(ITWorld, 2012)
- 예를 들어, 페이스북과 같은 소셜 네트워크 서비스의 경우 매달 300억개의 새로운 콘텐츠가 발생하고 이러한 거대한 데이터를 관리하기 위한 하드웨어 및 소프트웨어 등 모든 프로세스가 바로 빅데이터 플랫폼을 의미함
- 빅데이터는 통상적으로 사용되는 데이터 수집 및 관리, 처리와 관련된 소프트웨어의 수용 한계를 넘어서는 크기의 데이터를 말하며, 빅데이터의 사이즈는 단일 데이터 집합의 크기가 수십 테라바이트에서 수페타바이트에 이르며, 그 크기가 끊임없이 변화하는 것이 특징임(위키피디아 2013)
- 메타그룹(현 Gartner, Inc. 미국 정보기술연구 및 자문회사)의 애널리스트 더그 레이니(Doug Laney)는 2001년 그의 연구보고서 “3D Data Management : Controlling Data Volume, Velocity and Variety”와 관련 강의에서 데이터의 급성장에 따른 이슈와 기회를 “3V” 즉, 데이터의 양(volume), 데이터 입출력의 속도(velocity), 데이터 종류의 다양성(variety)이라는 세 개의 차원으로 정의함.
- 또한, 최근 2012년 가트너(Gartner, Inc.)는 “빅데이터는 큰 용량, 빠른 속도, 그리고 높은 다양성을 갖는 정보 자산으로서 이를 통해 의사결정 및 통찰발견, 프로세스 최적화를 향상시키기 위해서는 새로운 형태의 처리방식이 필요하다”며,

IBM의 경우 진실성(Veracity)을, 브라이언 휙킨스(Brian Hopkins)의 경우 가변성(Variability)을 추가하여 새롭게 “4V”를 정의함.

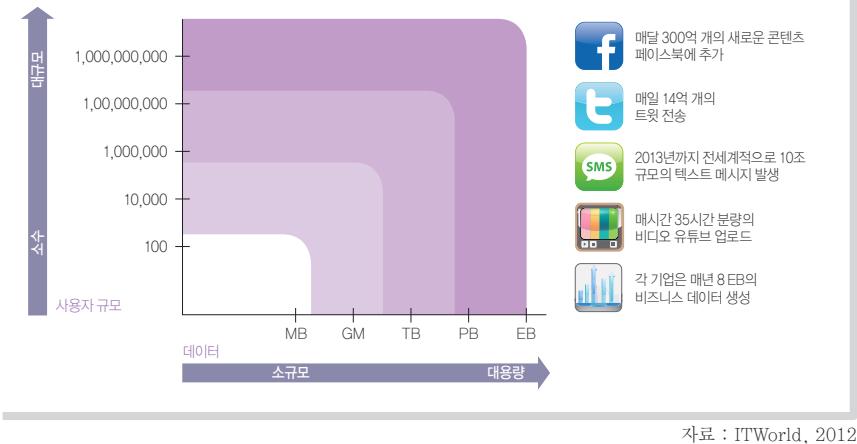
❷ 빅데이터의 활용을 위한 분석 기법은 크게 분석 기술과 표현 기술로 분류됨

- 빅데이터 분석 기술은 기존 통계학과 전산학에서 사용되던 데이터 마이닝, 기계 학습, 자연 언어 처리, 패턴 인식 등의 기술에 기반하고 있음
- 최근 비정형 데이터의 증가로 인해 자연 언어 처리 기술 기반 텍스트 마이닝 및 긍정과 부정을 의미하는 특정 언어 처리 기술 기반 오피니언 마이닝 등이 있으며, 행위자 간 연결구조 분석에 기반한 소셜 네트워크 분석 그리고 유사한 특징을 가진 집단 도출을 위한 군집분석 등이 특히 주목 받고 있음
- 빅데이터 표현 기술은 분석 기술을 통해 분석된 데이터의 의미와 가치를 시각화하기 위한 기술로서 NodeXL과 Gephi, R 기반 패키지 등이 있음

❸ 빅데이터를 활용한 국외 사례 및 의의

- 2008년 미국 대통령 선거에서 버락 오바마 후보는 다양한 형태의 유권자 데이터 베이스를 확보하여 이러한 빅데이터를 보트빌더(VoteBuilder.com) 시스템으로 통합 관리함으로서 기본 유권자 성향 분석 및 미결정 유권자의 의사 예측에 활용함
- 미국 온라인 종합 쇼핑몰 아마존닷컴(amazon.com)은 모든 고객들의 구매 내역을 빅데이터로서 수집, 관리 및 분석함으로서 고객 맞춤형 상품을 이메일과 홈페이지 상에서 중점적으로 제공하여 매출 제고
- 최근 구글 및 페이스북도 이용자의 기본 정보, 사진과 동영상 같은 비정형 데이터를 즉시 수집하여 분석함으로서 고객 맞춤형 광고를 제공하는 등 빅데이터 관련 기술의 적극적으로 도입하고 있음
- 빅데이터 관련 기술을 적극적으로 도입하고 있는 미국 의료부문의 경우, 연간 3,300억 달러(미 정부 의료 예산의 약 8%에 해당하는 규모)의 직·간접적인 비용 절감 효과를 보일 것으로 전망됨

그림1 ● 빅데이터 사용자 규모에 따른 데이터 규모의 증가

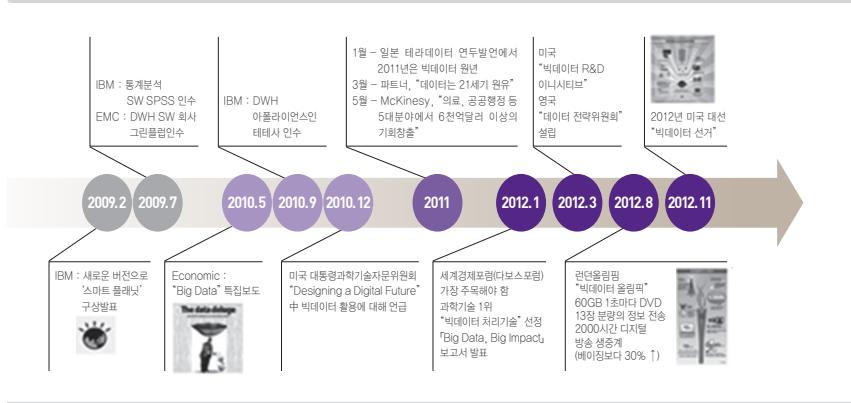


자료 : ITWorld, 2012

❸ 새로운 자원으로서의 빅데이터

- 빅데이터 자원 자체의 축적 및 관리뿐만 아니라 이를 가공하고 분석할 수 있는 차별화된 기술과 서비스가 바로 ‘개량된 자원’ 확보의 핵심
 - ‘개량된 자원’으로서 빅데이터는 공공기관, 기업체 및 시민들의 필요에 따라 구성 및 맞춤화 시킬 수 있는 미래의 동적 자원

그림2 • 새로운 자원으로서의 빅데이터 관련 글로벌 사례 추이

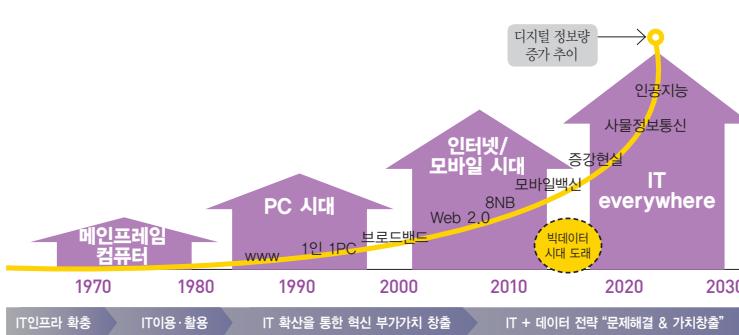


자료 : 한국정보화진흥원 2013

❶ 혁신과 새로운 산업 동력으로서의 빅데이터

- 혁신의 동력으로서 빅데이터는 기업 및 공공분야의 더 나은 서비스를 제공하는데 있어 반드시 필요한 기준을 제시
- 빅데이터에 대한 이해, 그리고 다양한 분야의 더 많은 연구개발과 투자를 통해 새로운 비즈니스 기회의 창출
- ‘개방된’ 빅데이터를 처리 및 분석할 수 있는 하드웨어/소프트웨어 인프라와 사용자 맞춤형 솔루션을 접할 수 있는 다양한 경로 확보

그림3 • IT 발전에 따른 데이터 변화 및 정부정책 패러다임 변화



자료 : 한국정보화진흥원, 2013

❷ 빅데이터 기반 사회를 향한 변화

- 빅데이터를 통한 공공 및 민간 부문의 변화는 비용 대비 효과성 및 정보 전달의 적시성 제고로부터 출발(Gartner 2012)
- 정보통신 기술의 발전으로 인한 빅데이터의 수집, 저장, 검색, 분석 관련 비용의 상대적 절감은 기존 경제학의 관념을 탈피한 혁신적인 비즈니스 모델 제공
- 특히 이러한 빅데이터 관련 기술의 발전은 빠르고 정확한 데이터의 수집과 분석을 통해 공공 및 민간 부분 내 조직들에게 필수적인 정보를 적시에 제공함으로써 비용 대비 효과성 제고에 기여
- 뿐만 아니라 빅데이터는 단지 데이터의 수집, 관리, 분석을 위한 기술적 변화를 넘어 정보의 개방과 공유 패러다임을 통해 다양한 이해관계자 간 효과적인 협력을 위한 최적의 의사결정 체계 제공

❶ 빅데이터 활용의 필요성

- 빅데이터는 행복하고 안전한 사회 구현을 위한 과거, 현재, 미래 분석 및 예측을 위한 자산
- 빅데이터는 혁신적인 아이디어와 다양성에 기반하고 있는 창조 경제 구현을 위한 미래 성장 동력
- 빅데이터의 개방을 통한 공공기관 및 기업체의 새로운 사업화 모델 촉진
- 빅데이터의 공유를 통한 정부 3.0 시대, 국민 맞춤형 서비스 제공

❷ 빅데이터의 활용을 위한 비전 및 전략

- 빅데이터 기반 Start-up 및 글로벌 강소 기업 육성을 위한 적극적인 지원
 - : 개발자를 위한 하드웨어 및 소프트웨어 인프라 제공, 우수 아이디어를 사업화하기 위한 창업경진대회 개최, 빅데이터 관련 기술 검증 및 표준화 지원
- 국내 빅데이터 전문 인력 양성을 위한 기반 제공
 - : 대학(원) 커리큘럼 및 전문강사 지원, 공공 및 민간 부문 등 빅데이터의 업무 활용을 위한 재직자 전문 교육과정 개설, 빅데이터 관련 우수 인력 발굴을 위한 분석 경진대회 개최
- 공공 및 민간 대상 빅데이터 기반 사업화 컨설팅 확대
 - : 빅데이터 도입이 시급한 공공 및 민간 조직을 대상으로한 사업화 방안 수립, 빅데이터의 활용을 위한 역량 진단을 통해 빅데이터 도입 가이드라인 신설, 빅데이터 관련 애로사항 해결을 위한 맞춤형 기술 컨설팅 수행
- 글로벌, 공공, 민간 빅데이터의 개방 및 공유를 통한 국가 미래전략 수립 지원
 - : 2014년 세부사업계획 수립 및 2015년 시스템 고도화를 통해 국가사회 현안을 해결하고, 미래 핵심 아젠다 및 중장기 국가 미래전략 개발
- 활용 가치가 높은 공공 및 민간 데이터를 지속적으로 확충하여 산업 활성화 및 전문 인력 양성을 위한 빅데이터 뱅크 구축
 - : 2014년 기관 간 협력, 공공데이터 개방(Open API) 및 시범사업 자료 확보를 통한 통합형 빅데이터 분석활용센터 개설

02

빅데이터 사례집의 목적 및 구성

❶ 빅데이터 사례집의 목적

본 사례집의 목적은 공공과 민간의 빅데이터사업에 모범이 될만한 선진사례(best practice)를 발굴하여 빅데이터 사업의 활성화와 산업발전에 기여

❷ 빅데이터 사례집의 구성

- 1장에서는 새로운 자원 및 미래 혁신과 새로운 산업 동력으로써의 빅데이터 개념과 이를 활용한 창조경제 지원 및 국민행복 구현을 위한 비전과 전략을 제시
- 2장에서는 2013년 미래창조과학부와 한국정보화진흥원이 빅데이터 도입 및 활용의 확산을 위해 선도적으로 추진한 6개 시범사업을 소개
- 3장에서는 광대역망 기반 농어촌 IT 융합 사과작황정보 서비스, 기상 빅데이터 분석을 통한 효율적인 위험기상 예측 서비스, 통합형 재난–안전 정보 네트워크 (스마트 빅 보드) 등 공공부문 사례 11개를 제시
- 4장에서는 스마트 통합물류 시스템, 맞춤형 여행 컨설팅 서비스 등 민간부문 사례 13개를 제시

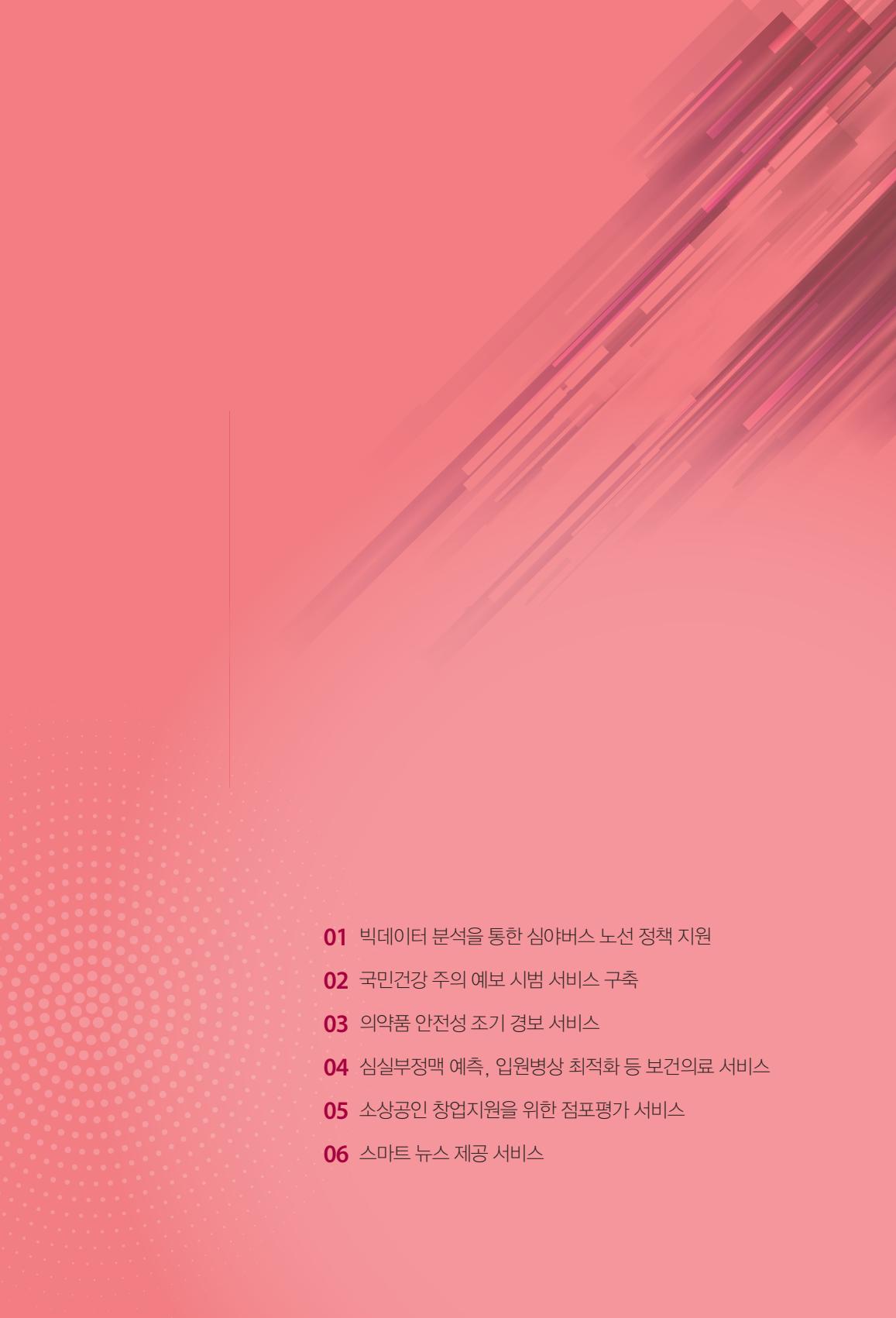
II

빅데이터 활용 사례 Part 1.

미래창조과학부 시범사업 6종

2013 미래창조과학부 빅데이터 시범 서비스

2013년 빅데이터 시범 서비스는 미래창조과학부와 한국정보화진흥원이 빅데이터 서비스를 공공·민간 영역에 적극 도입·확산하기 위한 목적으로 추진하였다. 공공 분야의 풍부한 데이터와 민간데이터를 연계·활용한 공익형 서비스로 수혜자가 체감할 수 있는 차별화된 우수 서비스 발굴을 통해 빅데이터 이용 활성화 및 기관 간 협력 활성화를 도모하였다. 매칭펀드 형태로 총 6개의 서비스가 개발되었으며 2013년 6월부터 12월까지 7개월간 진행되었다

- 
- 01** 빅데이터 분석을 통한 심야버스 노선 정책 지원
 - 02** 국민건강 주의 예보 시범 서비스 구축
 - 03** 의약품 안전성 조기 경보 서비스
 - 04** 심실부정맥 예측, 입원병상 최적화 등 보건의료 서비스
 - 05** 소상공인 창업지원을 위한 점포평가 서비스
 - 06** 스마트 뉴스 제공 서비스

01

빅데이터 분석을 통한 심야버스 노선 정책 지원

- 주관기관 : (주)KT
- 참여기관 : 서울특별시

야근이 잦은 직장인 A씨는 요근래 한시를 덜었다. 12시 넘어 끝나는 일이 많아보니 몸도 피곤하지만 택시비가 늘 부담이었고, 무엇보다 택시를 잡는 것 자체가 큰일이었다. 그러나 심야버스로 모든 문제가 해결됐다. 새벽 5시까지 다니기 때문에 대중교통의 공백이 사실상 없어졌다. 사람 많은 동네를 두고 먼 곳으로만 다니던 버스가 신기하게도 이번엔 동네로 들어왔다. 통신사의 유동인구 빅데이터를 분석한 과학적 노선 수립 덕분이다.

※ 서울시 심야버스 7개 노선('13.9월) 수립 지원

가. 추진 배경

- ❶ 심야시간대 택시 승차거부로 인한 불편, 심야 근로자의 교통수단 부재, 교통비 부담 등으로 인해 서울시는 심야시간대 대중교통의 편의 마련이 필요
- ❷ 수익성이 낮은만큼 최소의 비용으로 서울시내 전역을 지원할 수 있는 합리적 노선 수립방법이 필요한 상황이었고, 이를 위해 통신사업자인 KT와 협력하여 데이터 기반의 객관적 자료를 근거로 한 심야버스 노선 수립을 추진하게 됨

나. 추진 내용

❶ 활용데이터

기관명	데이터명
(주)KT	CDR(Call Detail Record), 고객통계 정보
서울시	시내버스 현황, 정류소 현황 등

▣ 주요 개발 내용

(주)KT의 CDR 데이터 약 30억건과 서울시의 택시이용 데이터 등을 연계 활용

- 지역별 유동인구 파악을 위해 서울시를 블록 단위로 분할하고, 목적지, 통행량 등을 산출하기 위해 거리 기반 알고리즘 제작
- 심야시간대 유동인구의 밀집도를 분석하여, 시민들의 이동경로에 있어 버스노선이 가장 필요한 곳을 중심으로 구역을 설정하고, 이용객의 최적화된 수준을 효율성의 관점에서 고려하여 배차간격 조정
- ※ (주)KT의 빅데이터 분석 플랫폼(KDAP) 활용
분석 결과를 활용하여 심야버스 5개 노선 검증, 수정 및 배차간격 조정

그림4 • 버스노선 수립을 위한 빅데이터 분석 도입 효과



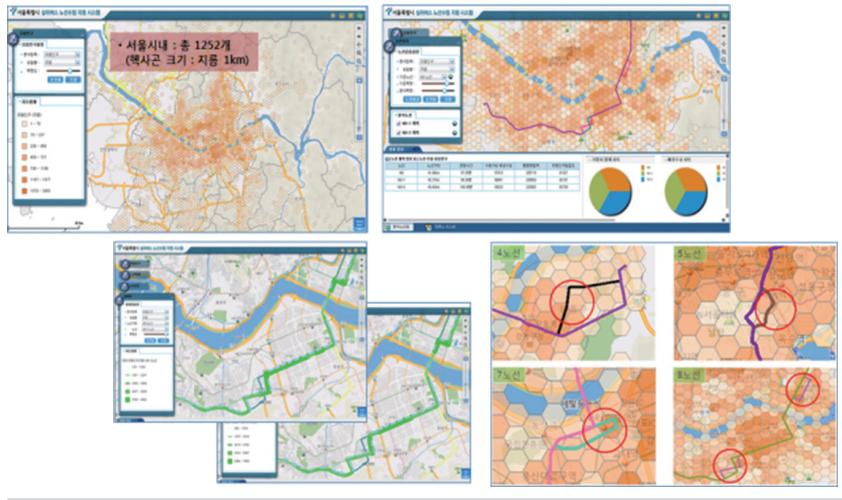
자료 : 한국정보화진흥원, 2013

그림 5 • 시스템구성도



자료 : 한국정보화진흥원, 2013

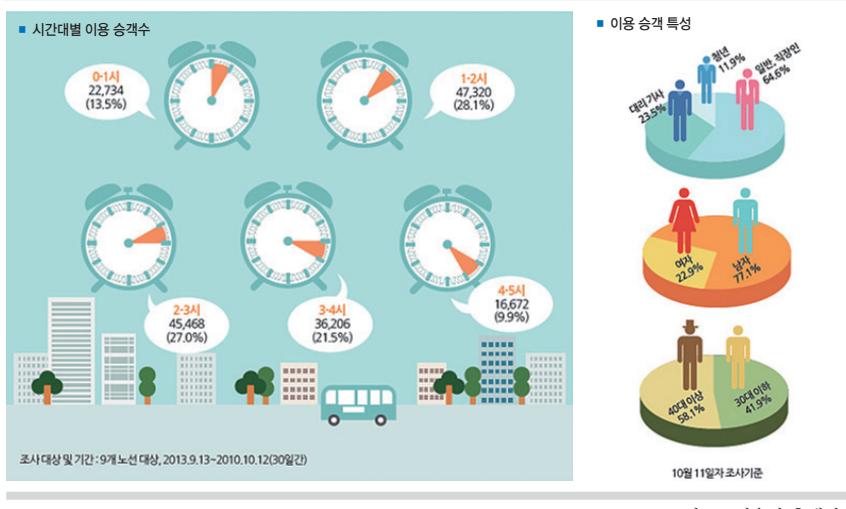
그림 6 • 유동인구 분석 및 적용과정



자료 : 한국정보화진흥원, 2013

- 빅데이터에 기반한 유동인구 분석은 시간대별 이용객 수뿐만 아니라 이용승객의 특성에 이르기까지 복합적이고, 구체적인 분석을 가능하게 함

그림 7 • 심야 전용 버스 이용 승객수 및 특성



자료 : 서울시 홈페이지

- 2013년 4월부터 실시된 ‘심야전용버스’ 정책지원 서비스는 시민들의 큰 호응에 힘입어, 심야시간대 유동 인구가 많은 지역을 중심으로 7개 노선이 추가되었고, 9월부터는 총 9개의 올빼미 버스노선이 운용 중에 있음

다. 성과 및 시사점

- 민원, 버스사업자 의견 등에 의해 결정되던 버스노선을 데이터 분석을 통해 결정함으로써, 데이터에 기반한 신뢰성 있고 과학적인 행정구현 및 시민편의 증진
- 버스 운송량 증가 및 버스 노선에 대한 지역 간 분쟁 소지를 차단하였고, 기존 시내버스에 비해 만족도 향상
※ 기존 시내버스 만족도 74.30점, 본 심야버스 만족도 80.15점
- 빅데이터가 앞으로 정책결정에 상당한 영향을 미칠 수 있음을 증명하였으며, 지방행정정보화 연찬회에서 본 서비스로 서울시가 대통령상 수상(2013.9.12)
- 빅데이터 융합, 분석 결과를 정책에 반영하는 획기적인 사례로 창조적인 행정 실현을 위한 민관 협력의 계기가 됨

02

국민건강 주의 예보 시범
서비스 구축

- 주관기관 : 국민건강보험공단
- 참여기관 : 다음소프트

어린이집을 운영하는 C씨는 아이들 중 한명이라도 열이 오르면 가슴이 덜컥 내려앉는다. 지난 여름에는 눈병이 유행하여 수많은 엄마들의 원망과 지탄에 시달려야 했다. 결국 비위생적인 어린이집이라는 오명 속에 몇 명의 아이까지 떠나갔다. 그러나 감염병 주의예보 서비스를 알게 되면서 걱정을 덜었다. 유행이 한차례 쓸고 지나간 뒤 대처할 수 밖에 없던 전과 달리, 2주전 미리 예보를 접함으로써 미리미리 손씻기, 소독 등으로 대비할 수 있게 되었기 때문이다. 엄마들한테도 미리 예보내용을 공지하여 주의토록 할 생각이다. 어린이집 이미지도 좋아질 것이다.

※ 국민건강보험공단 ‘건강인’ 웹사이트(<http://hi.nhis.or.kr>)에서 제공 예정

가. 추진배경

- ❶ 전염병 유행에 대해 기존 체계는 사후 대처만 가능하여 사회경제적 비용이 발생 할 수 밖에 없음
- ❷ 빅데이터 분석을 활용하여 예방과 예측 중심의 의료서비스 제공을 시도하고자 함
- ❸ 기상청의 일기예보와 같이 각종 전염성 질환의 위험도를 사전예보할 수 있는 체계를 구축하고자 함

나. 추진내용

❶ 활용 데이터

기관명	데이터명
국민건강보험공단	국민건강정보(진료내역 데이터)
다음소프트	SNS(트위터, 다음 카페) 데이터

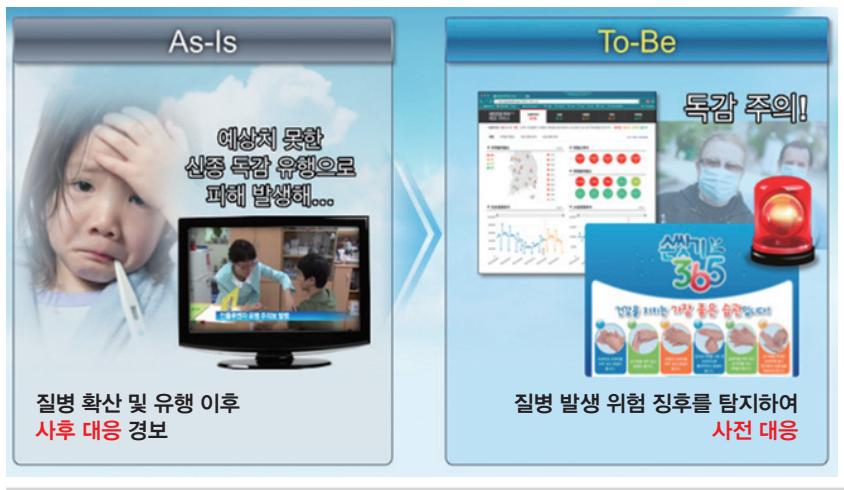
❷ 주요 개발 내용

- 현재 시점의 주요 유행성 질병에 대한 위험도 예보 및 지역별·연령별 위험도 예보 등 감염병 유행 징후 시 주의 예보 제공
- 독감을 예로 들면, 어떤 지역을 중심으로 환자가 늘고 있는지, 환자들의 연령대는 어떤지, 진료 환자 숫자가 얼마나 늘어나는지, 소셜미디어에서 검색하는 사람들의 주이는 어떤지 등을 파악해 미리 알려주는 방식
- 주요 유행성 질병에 대한 건강보험정보와 소셜미디어 정보를 융합하여 질병 예측 모델을 개발
- 주요 유행성 질병의 위험도와 동향을 한눈에 파악할 수 있는 대시보드 서비스를 제공하고, 지역별 주요 유행성 질병 위험도 정보, 지역 내 질병 관련 진료 동향 및 연령별 진료 현황정보 제공

❸ 주요 유행성 질병 진료현황, 과거 진료통계, 질병 동향 및 고위험 지역 정보 제공

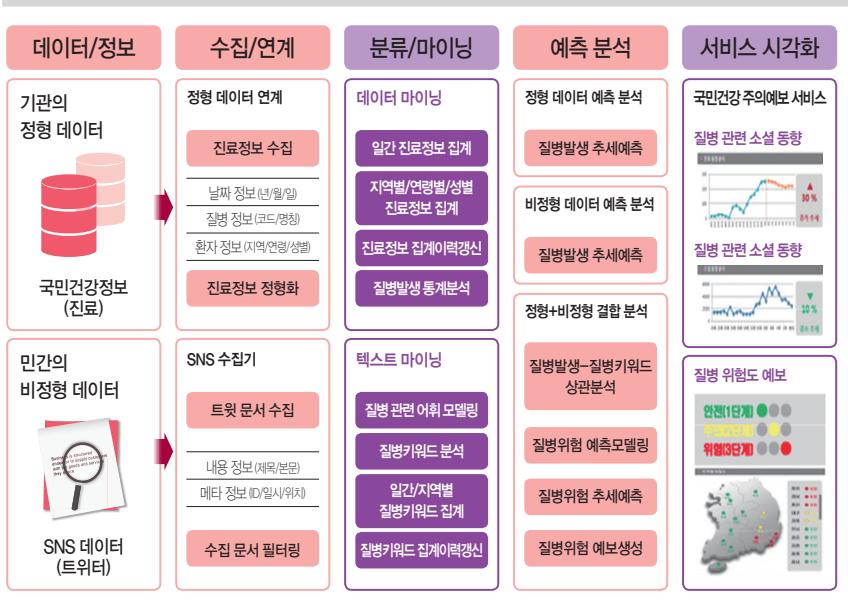
❹ 자연어처리기술을 이용한 텍스트마이닝 기법, 통계분석/기계학습 기술을 이용한 데이터마이닝 기법 등을 활용함

그림 8 • 국민건강 주의 예보를 위한 빅데이터 분석 도입 효과



자료 : 한국정보화진흥원, 2013

그림 9 • 국민건강 주의 예보 시범 서비스 과정



자료 : 국민건강보험공단, 2013

그림 10 • 국민건강 주의 예보 시범 서비스 시스템 구성도



자료 : 한국정보화진흥원, 2013

다. 성과 및 시사점

- 현재까지의 건강보험 진료서비스는 질병발생 후의 치료 중심이었으나, 국민 건강 주의예보 서비스는 급성기(전염병) 질병, 계절적·주기적 발생 질병, 다수 및 다빈도 질병 등 건강이상 징후에 대하여 사전에 건강주의 예보를 발동하는 것으로 앞으로 건강보험의 예방 중심의 서비스로 전환하는 커다란 계기가 될 것
- 주요 감염병 유행을 예측하여 관련 기관과 국민 개인이 유행에 대비하고 예방 할 수 있도록 지원함으로써 국민건강 증진 및 사회적 편익 극대화
- 공단은 건강주의 예보가 발동되면 지역별·집단별(가족, 학교, 직장 등) 맞춤형 건강관리 안내사업을 실시하고, 향후에는 희망자에게 개인별 맞춤형 모바일 건강서비스를 제공할 예정이어서 국민건강증진 및 사회적 편익이 극대화될 것으로 기대
- “공단이 보유한 국민건강정보 빅데이터를 활용해 질병예방과 생애주기별 맞춤형 건강서비스를 제공할 수 있도록 노력할 계획” (건강보험공단, 2013)으로 이는 공단 빅데이터의 경제적 사회적 가치를 높이는 주요 기반이 될 것

03

의약품 안전성 조기 경보 서비스

- 주관기관 : 에스지에이 주식회사
- 참여기관 : 한국의약품안전관리원, (주)와이즈넛

약사인 D씨는 요즘 많이 처방하는 ‘ㄱ’이라는 약에 대해 의문이 들었다. 효과는 좋은데 간혹 급성설사 증세를 보인 환자들이 있기 때문이다. 직감적으로 부작용이라는 느낌이 왔지만, 바로 대응은 망설여졌다. 잘 팔리는 제품인 이 약에 대한 부작용은 제약사에 영향을 줄 수 있기도 하지만, 내게 온 환자들에게서만 나타난 우연일 가능성도 배제할 수 없기 때문이다. 한 사람이라도 피해를 줄이기 위해 신고를 해야 하는지, 아니면 확실해질 때 까지 기다리는게 맞는지 혼란스럽다. 그러나 앞으로 의약계 종사자들은 이런 고민을 할 필요가 없게 될 전망이다. 의약품 부작용 조기경보 서비스가 도입되면 시스템이 보건·의료 모든 데이터를 모니터링하며 자동으로 위해사례를 검출해내어 경고를 보내게 될 것이다.

※ 한국의약품안전관리원이 서비스 모델 개발 및 검증 추진 중

가. 추진배경

- ❶ 의약품 부작용이나 오남용에 대한 조치는 자발적 신고에 의지하고 있어, 대부분 사례가 만연된 후 사후 조치 위주로 진행되고 있음
- ❷ 빅데이터 분석을 통한 의약품 부작용이나 오남용 사례의 조기 검출 및 빠른 조치를 통해 이로 인한 피해와 사회 경제적 비용 손실을 최소화하고자 함

나. 추진내용

❶ 활용데이터

기관명	데이터명
한국의약품안전관리원	유해사례신고DB
와이즈넛	SNS(블로그, 지식인, 카페, 트위터), 뉴스(웹), 문헌(웹) 데이터
아주대학교병원	EMR(전자의무기록)

❷ 주요 개발 내용

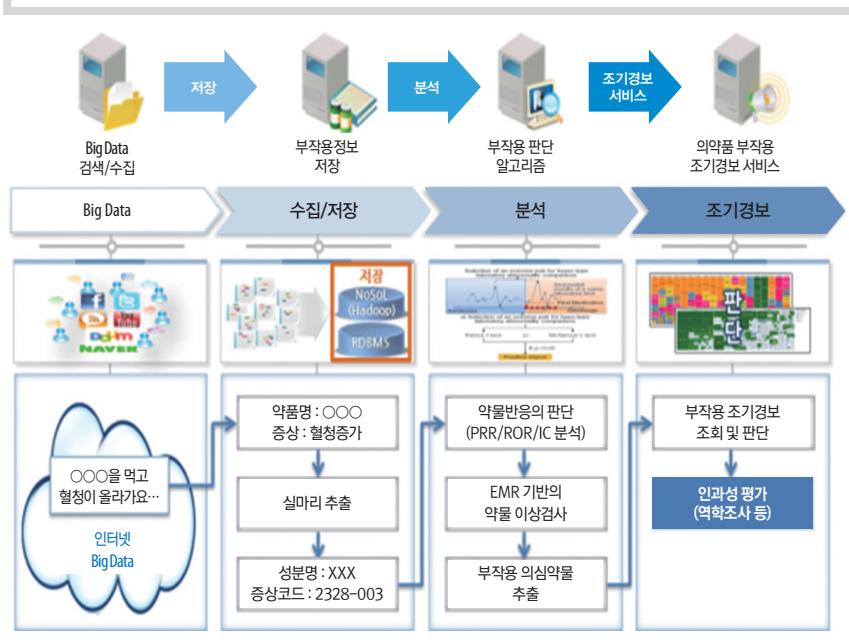
- 한국의약품안전관리원이 보유한 유해사례 신고데이터와 인터넷 포털, 소셜 사이트 등의 빅데이터에서 의약품 부작용, 의심정보 등을 수집·분석하여 유의 의약품에 대해 조기경보
- 빅데이터에서 의약품 부작용 및 오남용 사례를 수집 및 분석
- 의약품 부작용 검증을 위해 추출된 실마리정보를 제공
- 의약품 부작용에 대한 검증 결과, 의약품 오남용 사례 및 분석결과 제공
- 의약품 부작용 가능성을 병원 의무정보를 기반으로 확인하여 조기 인지하고, 조기 대응을 위한 정보 제공
- 자연어처리기술을 이용한 텍스트마이닝 기법 사용, 부작용 검증을 위해 코호트 기반 연구방법, 환자/대조군 비교 알고리즘 등 사용

그림 11 • 의약품 안전성 조기 경보를 위한 빅데이터 분석 도입 효과



자료 : 한국정보화진흥원, 2013

그림 12 • 의약품 안전성 조기 경보 서비스 프로세스



다. 성과 및 시사점

- ❶ 국민복지와 건강을 위협하는 의약품 부작용 및 오남용을 조기에 발견하여, 선제적 대응
- ❷ 우리나라에 적합한 약물부작용 조기경보시스템 개발을 통해 빅데이터를 활용한 의료 약물 부작용 조기경보 기술 확보 및 유용성 검증
- ❸ 의약 정보와 규정에 기반한 분석으로 추출된 데이터에서 실제 부작용 가능성을 판단하고, 향후 의심되는 의약품에 대한 정보를 관련 기관 및 일반 소비자에 공유하는 시스템 구축 예정

04

심실부정맥 예측, 입원병상
최적화 등 보건의료 서비스

- 주관기관 : 서울아산병원
- 참여기관 : 한국전자통신연구원, 한국마이크로소프트(유), (주)테크아이, 켐아이넷(주), 한국빠스소프트웨어(유)

외과의사인 E씨에게 심실부정맥은 답이 없다. 발생하면 사망에 이르는 경우가 많다. 심지어는 중환자실에 입원 중인데도 손을 놓쓰는 경우가 있다. ‘몇분 아니 몇초 전에만 알아도 사망은 막을 수 있었을텐데.’, ‘부정맥 발생은 신만이 미리 알 수 있는 것인지.’ 그런데 이제 의사도 미리 알 수 있게 될 전망이다. 센서 데이터를 분석하여 적어도 10초전에 경고를 보내주는 서비스가 빅데이터 분석 기술을 토대로 개발되었기 때문이다.

가. 추진배경

- ❶ 기존에 이루어진 환자 모니터링은 단순한 정보수집에 지나지 않아 그 활용정도가 제한되었으나, 빅데이터 활용을 통한 국민건강 증진서비스에 대한 요구증대

나. 추진내용

- ❷ 활용데이터

기관명	데이터명
서울아산병원	실시간 환자 심전도/심박수, 진료 데이터, 입/퇴원 기록, 병원 경영자료
다음	검색 데이터, SNS 데이터
국립과학수사연구원	국과수 마약류 관련 DB

- ❸ 주요 개발 내용

- 인플루엔자 유행 예측 서비스 : 인터넷 포털사이트의 검색어 데이터와 SNS 데이터를 분석하여 인플루엔자 현황 및 예측 정보를 웹서비스 형태로 제공

- 심실부정맥 예측 서비스 : 중환자로부터 수집되는 심실 모니터링 데이터를 빅데이터 플랫폼에 저장하고, 패턴 분석을 통해 부정맥 발생상황을 사전 예측

그림 13 • 심실부정맥 예측을 위한 빅데이터 분석 도입 효과



자료 : 한국정보화진흥원, 2013

그림 14 • 심실부정맥 예측 프로세스



자료 : 한국정보화진흥원, 2013

- 입원 병상 배정 최적화 서비스 : 병원이 보유하고 있는 환자 입원/진료 데이터, 입퇴원 기록, 보험 청구 데이터 등 병원 경영자료를 복합 분석하여 각 과별 병상 배정을 최적화
- 계절적 요인, 병원 의료진의 수, 진료 행위의 변화를 모델에 반영하며 이를 토대로 준 실시간(semi-real time) 입원 병실의 가동률의 변화 패턴을 파악]
- 환자들의 입원 대기시간, 장기재원일의 감소를 통해 조기 치료 가능

그림 15 • 입원 병상 배정 최적화를 위한 빅데이터 분석 도입 효과



자료 : 한국정보화진흥원, 2013

그림 16 • 입원 병상 배정 최적화 프로세스



자료 : 한국정보화진흥원, 2013

- 마약류 인지 · 감시 서비스 : 마약류 유통 차단을 위해 SNS 데이터와 국립과학수사 연구소가 보유하고 있는 자체 DB를 연계 분석함으로써 신종 마약 발생 조기 감지 및 대처

다. 성과 및 시사점

- 심실빈맥이나 심실세동 상황으로 진입하기 전에 의료진에게 경보를 알려주어 중환자실 환자들의 생존율 향상 기대
- 다양한 의료 기기들에서 자동으로 기록되고 있는 의료장비 로그 데이터 분석을 위한 기반기술 확보
- 병원 경영측면에서 제한된 자원의 효율적 분배를 통해 각 임상과별로 균형적, 안정적인 병상 가동률 유지
- 의료기관의 자원을 효율적으로 활용하여 국민보건의료비용 절감

05

소상공인 창업지원을 위한 점포평가 서비스

- 주관기관 : (주)오픈메이트
- 참여기관 : 비씨카드(주), 한국감정원

최근 명예퇴직한 B씨는 카페를 창업할 계획이다. 퇴직금을 모두 투자해야 하기 때문에 신중할 수밖에 없으나, 마땅히 창업정보를 얻을 길이 없다. 프랜차이즈 설명회에도 가보고 부동산도 찾아가 봤지만 신뢰해도 되는지 확신이 서질 않는다. 그러나 스토어스토리* 서비스를 알고 나서 눈이 번쩍 띠었다. 서울 시내 어디든 건물 내 점포의 개폐업 이력, 매출정보와 데이터 기반의 과학적 상권분석, 적정업종, 추정매출 등 정말 알고 싶었던 정보를 얻을 수 있었기 때문이다.

* <http://partners.bccard.com>에서 시범 서비스 제공

가. 추진배경

- ❶ 현재 국내 자영업 비율은 OECD 평균의 두배가 넘는 반면, 3년 내 폐업율은 50%가 넘어 매우 치열한 경쟁 속에 놓여있음
- ❷ 창업 준비과정에서 개인 및 영세 자영업자는 프랜차이즈 설명회, 부동산 사업자의 컨설팅 등에 의존할 수 밖에 없는 등 창업정보 인프라가 매우 열악한 실정
- ❸ 민간, 공공의 데이터를 연계 활용하여 소상공인의 창업 의사결정에 도움이 되는 정보 인프라를 제공하고자 함

나. 추진내용

● 활용데이터

기관명	데이터명
비씨카드	월 2억건씩 누적되는 카드거래데이터 중 서울지역 2억건(3년치)
한국감정원	부동산 임대시세 및 건축물 대장 데이터
소상공인진흥원	소상공인상가정보(점포명, 업종, 주소, 전화번호) 총 1억건(3년치)
공공 오픈데이터	인구, 가구세대, 36만 블럭

● 주요 개발 내용

- 약 1억건의 상가업소데이터, 6억건 이상의 카드사용 트래픽 데이터 기반 분석
- 동일지점의 최근 3년간 점포 개폐업 이력 추적
- 점포별 수년간의 개폐업이력을 추적하고, 업종변화별로 해당 점포의 매출 추정
- 창업자가 어떤 업종으로 창업하면 가장 높은 매출을 낼 수 있을지, 어떤 업종일 때 영업기간이 짧고 폐업율이 높았는지, 점포의 입지는 어떤 수준인지 등 창업결정을 위한 지표정보 제공
- 임대시세, 추정매출, 점포진단평점 등 입지상권분석에 필요한 기초정보 제공

그림 17 • 점포평가 서비스를 위한 빅데이터 분석 도입 효과



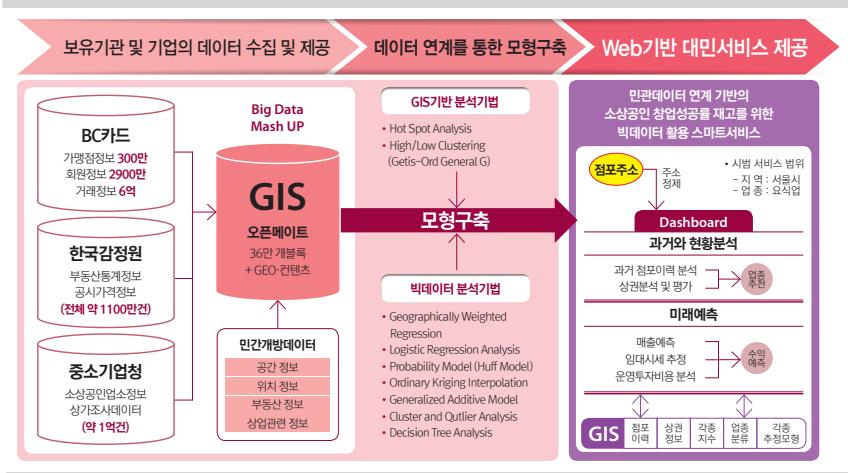
자료 : 한국정보화진흥원, 2013

그림 18 • 창업자의 고민과 현 수준

	사례	필요성
창업자의 고민	<ul style="list-style-type: none"> - 내가 창업하고자하는 업종이 현 위치에서 과거 실패한 경우가 있는가? - 이 자리에서 가장 높은 매출을 일으킬 만한 업종은 무엇인가? - 임대료, 권리금은 적정한가? - 예상 매출에 비해 비용이 높지 않은가? 	<ul style="list-style-type: none"> 부동산 중개업자 제공 정보 프랜차이즈컨설턴트 분석 의견, 창업전문가 분석 의견에 의존한 창업 → 점포거래, 가맹로열티의 이해관계로 인하여 객관적이고 정확한 정보제공이 되지 않음
정부의 지표	<ul style="list-style-type: none"> - 1년에 약 36%의 손비꿈이 일어나는 시장 (통계청 생활을 발표 2011년, 국세청 2011년 829,669개 개인사업자 폐업) - 국내 소상공인수 (265만, 소상공인진흥원) (통계청 성명을 발표 2011년, 국세청 2011년 829,669개 개인사업자 폐업) - 베이비부머 세대의 본격적인 은퇴시기 도래 (1955년 ~ 63년생 695만 명 : 통계청) 	<ul style="list-style-type: none"> 다산, 다사 구조의 창업시장 억순환 신중한 진입을 위한 정보서비스 미비 → 창업시장 경쟁구조 변화 필요
관련 서비스	<ul style="list-style-type: none"> - 소상공인진흥원 상권정보서비스 상가업소의 현황만을 제시 : 예측 서비스 미비 - 민간서비스 (SKT지오비전, 나이스비즈앱, 창업포털) 현상파악 위주 insight 제시 부족, 유료모형, 점포증개 등 수익 주구 - 국가 데이터 공개 정책 확대 데이터 공개 후 활용에 대한 컨텐츠 확보 필요 	<ul style="list-style-type: none"> 현황정보만을 제공하여 의사결정 지원이 어려움 → 빅데이터 분석을 통한 insight 발굴 미래예측 정보 제공이 필요

자료 : (주)오픈메이트, 2013

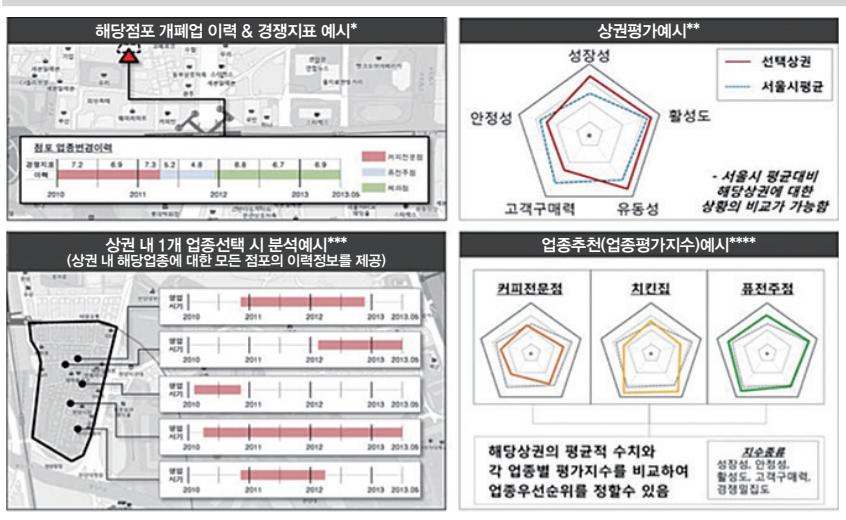
그림 19 • 서비스 플랫폼 구성도



자료 : (주)오픈메이트, 2013

- 점포의 과거 개폐업 이력과 더불어 다양한 경쟁지표를 중심으로 상권을 평가하여 특정업종에 대한 평가와 더불어 업종 추천도 가능하게 함

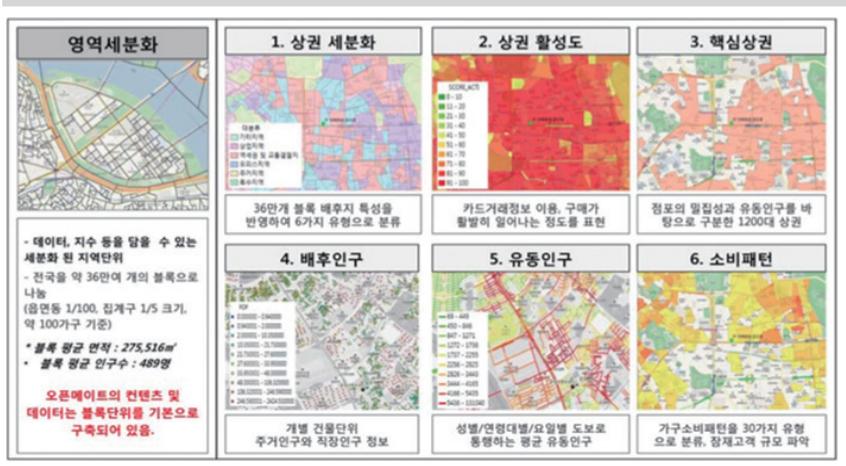
그림 20 • 상권 평가 및 업종 추천



자료 : (주)오픈메이트, 2013

- 행정동 단위의 통계데이터를 보다 활용도 높은 36만개의 블록으로 재가공하고
- 이에 대해 다양한 컨텐츠를 만들어 상권과 입지분석에 활용할 수 있도록 제공

그림 21 • 점포 이력 서비스 활용



자료 : (주)오픈메이트, 2013

다. 성과 및 시사점

- ❷ 유동인구가 많은 지역을 찾아내고 먹거리 · 의류 등 상권별 특성을 과학적으로 파악할 수 있는 지도가 마련되면 자영업자들의 실패 확률을 줄일 것으로 기대
- ❸ 성공률 높은 업종, 상권 조건이 우수한 점포를 기반으로 선별창업을 할 수 있고, 부정확한 컨설팅으로 인한 피해나 비용지출 감소
- ❹ 창업 점포의 부동산 가치를 객관적으로 판단함으로써 창업 과정에서 나타나는 과잉지출 또는 피해를 최소화

06

스마트 뉴스 제공 서비스

- 주관기관 : 차세대융합기술연구원
- 참여기관 : 서울대학교 산학협력단, (주)이스플러스

회사에서 보도자료 스크랩 업무를 하는 F씨는 요즘 신문기사에 불만이 많다. 하루 만여개 이상 쏟아지는 기사 흥수 속에 중복된 기사도 수도 없이 많아, 필요한 핵심기사를 정확히 찾아내기란 여간 어려운게 아니다. 그러나 스마트 뉴스 서비스가 개발되어 기사를 단순 취합·요약만 하는 것이 아니라 시간별·매체별로 검색·요약하고, 기사 의미 관계를 기반으로 관련 기사를 제공하며, 뉴스정보원 연결망을 시각화해서 보여줌으로써 이슈에 영향력을 미치는 전문가를 식별하는 등 핵심기사 위주로 효율적으로 뉴스를 볼 수 있게 되었다.

※ <http://www.kinds.or.kr/>에서 제공 예정

가. 추진배경

- ❶ 보도기사의 저널리즘적 가치는 사실성, 다양성, 심층성을 핵심으로 하는데 반해, 낚시성, 선정성, 광고성, 편향성 등을 가진 보도자료들이 범람함에 따라 독자들이 고품질의 정제된 기사들만 접하기가 어려움
- ❷ 기존 뉴스 모바일 앱의 단순 취합하여 요약해서 보여주는 수준의 한계를 넘어 뉴스기사를 시간별, 매체별로 검색, 요약하며, 중요한 순서대로 기사를 보여주고 의미관계를 파악하게 도와주는 스마트 뉴스 서비스 필요

나. 추진내용

④ 활용 데이터

기관명	데이터명
한국언론진흥재단	카인즈 기사 아카이브 (66개 매체의 1990년 이후 기사)

⑤ 주요 개발 내용

- 다양한 매체에서 생산되는 뉴스기사들을 주제별로 수집하고, 텍스트 마이닝, 소셜 네트워크 분석, 빅데이터 처리를 통해 뉴스 요약 및 시각화 서비스를 제공하는 모바일 앱, 뉴스 앱 개발
- 자연어 처리기술을 이용한 텍스트마이닝 기법을 사용하여 기사 아카이브 분석
- SNA를 활용한 중요도 부여 및 관계 설정을 통해 뉴스정보원 연결망 분석
※ HTML5 등 시각화 도구를 이용하여 분석 결과 시각화
- 주요매체의 기사를 요약 · 종합하여, 중요도에 따라 보여주고, 주제를 묶어주고, 소속별 의견을 대조하는 스마트 뉴스 앱 개발
- 뉴스기사를 시간별, 매체별로 검색, 요약하고, 중요한 순서대로 기사를 보여주고, 의미 관계를 파악하게 도와주는 스마트 뉴스 서비스 구현
- 다양한 언론 매체에서 생산된 기사들을 수집하고, 주제별로 기사 내의 사실을 중심으로 문장 단위로 기사를 요약하여 보여줌
- 스마트 패드 또는 태블릿PC 등의 모바일 기기에 최적화된 뉴스기사 요약 서비스 제공
- 뉴스정보원연결망을 시각화하여 보여줌으로써 사회이슈에서 영향력을 발휘하는 주요한 전문가의 식별 및 관계 정보 파악

그림 22 • 스마트 뉴스 서비스를 위한 빅데이터 분석 도입 효과



자료 : 한국정보화진흥원, 2013

그림 23 • 스마트 뉴스 서비스 프로세스



자료 : 한국정보화진흥원, 2013

다. 성과 및 시사점 도출

- 디지털 뉴스 유통의 새로운 모형을 제안
- 사실성, 다양성, 심층성이라는 Journalism적 가치를 구현한 새로운 뉴스 서비스로서 언론 서비스 발전에 기여

참고문헌

국민건강보험공단 <http://www.nhic.or.kr>

국민건강보험공단(2013), ‘건강보험 빅데이터 현황 및 활용계획’

국민건강보험공단(2013.6.19), 건보공단, 일기예보와 같이 『국민건강 주의 예보』서비스 제공 추진

디지털 타임즈(2013.6.13.), SGA, 미래부 주관 ‘빅데이터 활용 스마트서비스 시범사업’ 수주

매일경제(2013.12.17.), 그 많은 정보를 분야별로 주루룩…스마트 뉴스 앱 뜬다

미래창조과학부(2013.12.18.), 빅데이터 페어 2013

백승민(2013), ‘빅데이터를 활용한 보건산업 신산업 전망 및 정책방향’. 한국산업보건진흥원. Vol. 84

서울시(2013), 2013 희망서울 시정운영 계획

송태민(2012), ‘보건복지 빅데이터의 효율적 활용 방안’. 보건복지포럼

스카이 데일리(2013.6.15.), 미래산업 ‘빅데이터’ 먼 훗날엔 재양일까

아시아 경제 (2013.6.13.), 미래부, 빅데이터 시범 서비스 개발 착수

중앙일보(2013.6.14.), SGA, 미래부 빅데이터 사업 선정..”SI 선두주자 발돋움”

중앙일보(2013.6.19.), 건보공단, 빅데이터 활용해 ‘건강 주의 예보’한다

파이낸셜 뉴스(2013.6.13.), SGA, 미래부 주관 ‘빅데이터 활용 스마트서비스 시범사업’ 수주

한국경제 DB인사이트(2013.6.18.), 미래부, 빅데이터 시범 서비스

한국정보화진흥원(2013.12.18.), 빅데이터 성과물 한자리에…빅데이터 페어 2013 개최

Enews.com(2013.4.16.), kt-서울시, 빅데이터 활용 ‘심야버스’, ‘와이파이 확대’

News Wire(2013.4.16.), kt, 서울시와 빅데이터 활용해 공공분야 복지증진에 협력

News Wire(2013.6.19.), 건보공단, ‘국민건강 주의 예보’ 서비스 제공 추진

NSP 통신(2013.4.16.), KT, 서울시와 빅데이터 활용 심야버스 노선정책 지원 등 공공서비스개선

SGA(2013.8.16.), SGA, 미래부 주관 ‘빅데이터 활용 스마트서비스 시범사업’ 수주

I. 빅데이터 시대

II. 빅데이터 활용 사례 Part 1.
미래창조과학부 사업사업 6종

III. 빅데이터 활용 사례 Part 2.
기본서적 11종

IV. 빅데이터 활용 사례 Part 3.
인간서적 13종

III

빅데이터 활용 사례 Part 2.

공공사례 11종

01 빅데이터 기반 영주사과작황정보 서비스

(경상북도, 미래창조과학부)

02 기상 빅데이터 분석을 통한 효율적인 위험기상 예측(기상청)

03 안전한 금융서비스를 위한 여신상시감시시스템(금융감독원)

04 특수질병 및 질환 예방 및 관리(보건복지부)

05 통합형 재난안전 정보 네트워크, 스마트 빅 보드(안전행정부)

06 맞춤형 정보 유통 서비스(안전행정부)

07 지리정보 통합한 지리적 프로파일링 시스템 구축(경찰청)

08 국가공간정보 통합데이터베이스 구축(국토교통부)

09 고용정보통합분석시스템 구축(한국고용정보원)

10 역외탈세 정보공유(관세청, 국세청)

11 빅데이터 분석을 통한 2013 국민 인식 변화 분석(문화체육관광부)

01

빅데이터 기반 영주사과작황정보 서비스

– 경상북도 · 미래창조과학부 · 한국정보화진흥원

가. 추진 배경

경상북도의 지역적 특성에 기반한 사과재배의 중요성

- 경상북도의 농업인구는 도민의 18%이고, 농업면적 또한 28만ha로 전체 가용 면적의 49%에 달함(경상북도 공공분야 빅데이터 활용방안 연구, 2013)
- 전국 사과 생산량의 63%가 경상북도에 의존하고 있는 상황이며(영남일보, 2012), 그 중에서도 영주는 전국 최대 사과 생산의 주산지로서 생산량이 7만여 톤에 이릅니다(매일신문, 2012)

기존 정부 사업의 결과로 광대역망이 구축되어 있음

- 2010년 4월, 방송통신위원회와 범부처, 지자체, 통신사업자 등이 참여하여 농어촌 스마트 인프라 구축 시작(한국정보화진흥원, 2010)

생산량 제고를 위한 기존 u-IT 시스템의 문제점 고찰

- 2012년부터 온도, 습도, 토양 내 수분, 강우 등과 같은 환경적 특성이 u-IT 기술을 통해 측정되는 생장환경관리 시스템이 영주시내 사과 재배 농가 중 13가구에서 실시 · 운용되고 있음
- 하지만, u-IT 센서 설치에도 불구하고, 기기관련 유지 · 보수가 적절하게 이루어 지지 않거나, 일부 기능이 미비하여 시스템 활용에 애로점이 상존함

❷ 사과 생산성 제고 및 적절한 사과 판매가격 책정에 기여

- 농업 관련 공공정보와 시민의 민간정보를 통합한 데이터를 분석하여 병충해 징후 발견, 생산량 추정, 가격동향 등의 사과작황정보를 제공하여 궁극적으로 사과 생산량을 증가시키고 적절한 사과가격을 책정하는데 참고자료로 활용 가능하도록 함

❸ 경북도 영주사과 빅데이터를 수집하고 정제한 후 텍스트마이닝과 데이터마이닝을 실시하고 빅데이터 모형을 개발해 검증해 봄으로써 우리나라 빅데이터 연구와 산업의 발전에 기여하고자 함

나. 추진 내용

❶ 빅데이터 셋 구축

- 정형 데이터

표 1 정형 데이터

항목	세부항목	수집기간	데이터량(건)
기후	기온, 습도, 일조, 풍속, 강수량, 기타	2000 ~ 2013년	90,608
u-IT	기온, 습도, 토양수분	2011년 11월 ~ 2013년	약 200만
병해충	복숭아순나방, 복숭아심식나방, 애모무늬잎말이나방, 기타	2005 ~ 2012년	2,058
생산량	전국 및 영주 사과생산량, 재배면적	2000 ~ 2012년	39
가격	전국 및 영주산 가격, 대체과일가격, 물가지수, 사과지수	2003 ~ 2013년	약 185만

- 비정형 데이터

표 2 비정형 데이터

항목	수집기간	데이터량(건)
영주, 사과, 가격, 병해충, 생산	2000 ~ 2013년	약 20만
관련 핵심어		약 70만

* 핵심어는 웹&SNS 기반으로 데이터를 수집한 데이터 다섯가지 분석단어(영주, 사과, 가격, 병해충, 생산)로 수집된 자료를 분석하여 생성한 단어.

빅데이터 모형 개발

- 사과 작황에 영향을 미치는 병충해 징후 발견과 사과 생산 추정 및 가격 동향 정보의 제공을 위하여 관련 데이터를 수집하고 정확한 서비스 모형 개발
- u-IT 센서와 기상청으로부터 수집한 기후 데이터와 페르몬 트랩을 통해 수집된 병해충 유찰 수 등 정형 데이터와 웹과 SNS의 병해충 데이터 등 비정형 데이터를 수집, 분석해 병해충의 발생 경위와 정보를 공유, 농민들에게 농약 살포 시기 등 유용한 정보를 제공

빅데이터 분석 결과 제공

- 2014년 1월부터는 스마트폰이나 컴퓨터, IPTV 등을 통해 관련 서비스를 제공할 계획이며, 농촌 광대역망 구축 및 활용을 위하여 마을회관 5곳에 IPTV, 사과재배 농가 25곳에 와이파이를 설치하여 결과 제공

사과작황정보서비스의 정의

● 사과작황정보

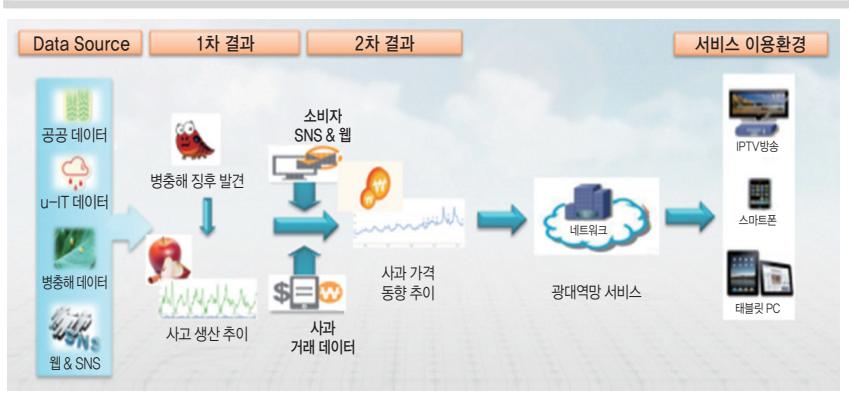
- 사과작황에 영향을 미치는 요소들인 농업, 기후, 병충해, 사과관련 웹 & SNS 등 데이터를 수집하여 병충해 및 사과 작황의 패턴 및 추이를 분석하고 관련 요소들의 상관성을 통해 예측할 수 있는 정보

● 사과작황정보 서비스

- 광대역망 기반의 IT 기기인 스마트폰, 태블릿 PC, IPTV를 통해 사과작황의 생산이나 품질에 영향을 미치는 다양한 데이터를 수집, 분석하여 병충해 징후 발견 정보 및 사과작황정보를 실시간으로 제공하는 서비스

❷ 사과 작황 정보 서비스 개념도

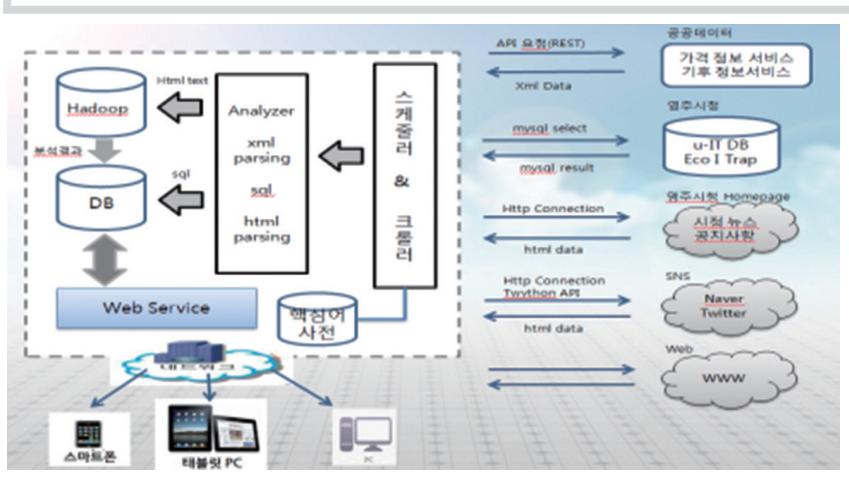
그림 24 • 사과 작황 정보 서비스 개념도



자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

❸ The SCRM¹⁾에 기반한 사과작황서비스 시스템 구조

그림 25 • The SCRM에 기반한 사과작황서비스 시스템 구조



자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

1) data Search, Curation, Refine, Matrix의 약자로 국내외 포털사이트와 트위터, 유튜브, 페이스북 등 다양한 미디어에서 데이터를 수집하여 원데이터(Raw Data)를 통합하여 제공해 주는 시스템(더아이엠씨)

❷ 빅데이터 기반 영주 사과 작황 정보 플랫폼

그림 26 ● 영주 사과 작황 정보 플랫폼 메인 화면

HOME | LOGIN | JOIN | SITEMAP | 경상북도 | 영주시

커뮤니티 | 서비스안내

농업 빅데이터를 수집하여
병해충 징후 및 사과 생산 추이,
가격동향 서비스를 제공합니다.

사과 공공데이터 | Apple Public Data

가격·증후 | 생산량·추이 | 병해충·징후

영주시 | 가격 | 생산 | 병해충

영주시 | 경상북도 | 영주시민 | 시과사랑증후원 | YeongJu | 한국정보화진흥원

개인정보 취급방법 | 이용약관 | 오바일

750-701 경상북도 영주시 시월로 5 (충현2동 470번지) 대표전화 054-634-3110 팩스번호 054-634-7770
COPYRIGHT © 2013 YEONGJU-CITY, ALL RIGHTS RESERVED.

자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

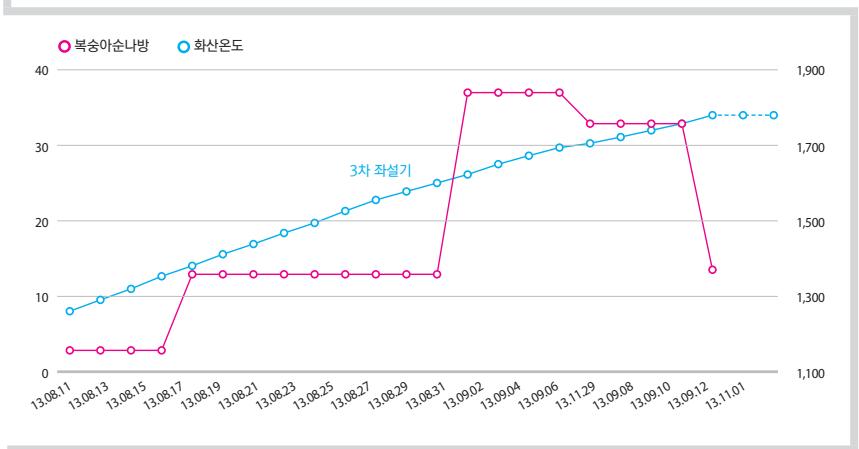
– 웹, 블로그, 카페, 트위터, 페이스북, 유튜브 등의 비정형 빅데이터(Unstructured Big Data)는 The SCRM을 통해 수집하고 처리한 후 패턴 분석과 상관성을 통한 예측분석을 실시하여 정보를 제공함

다. 성과 및 시사점

❶ 병충해 징후 발견

– u-IT 센서 데이터와 기상청의 기후데이터를 분석하여, 병충해의 징후를 발견함으로써 사전에 예방할 수 있도록 하여, 궁극적으로는 사과 생산 증대에 기여

그림 27 ● 김집 농가 복숭아순나방 발생 징후 그래프

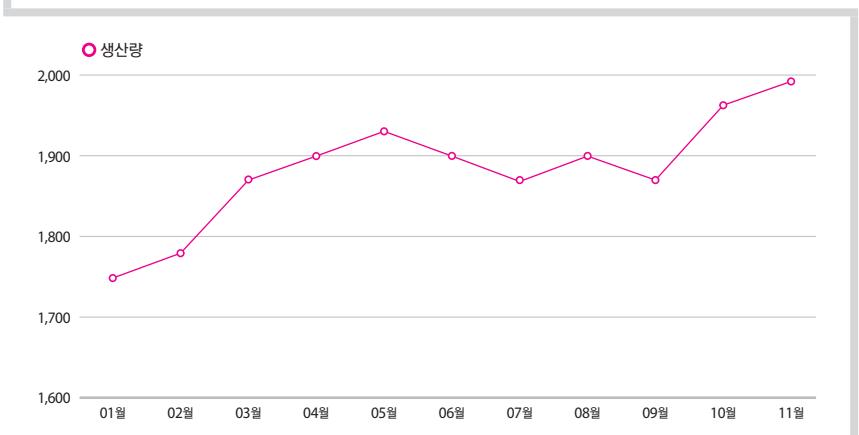


자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

❷ 생산 추정

- 사과생산 정형데이터와 웹 & SNS상의 사과 생산량과 관련된 비정형 데이터를 수집 및 분석하여, 앞으로의 생산량을 추정함으로써 사과 마케팅과 관련 정책 수립에 참고자료로 활용

그림 28 ● 영주시 사과 생산량 추이 그래프

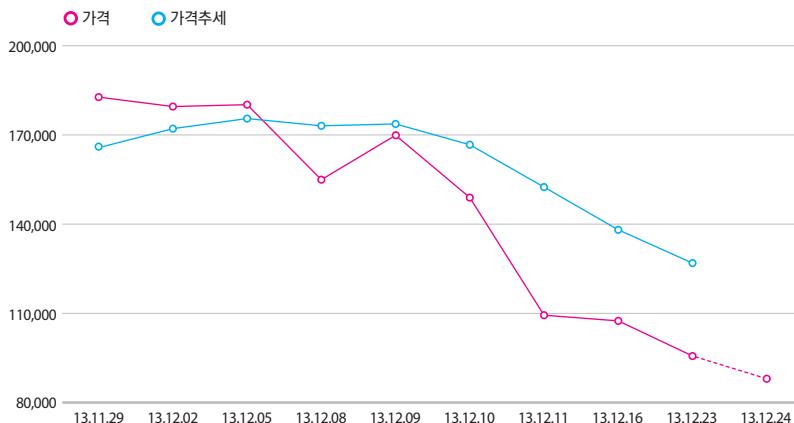


자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

❸ 가격동향

- 사과가격 정형데이터와 웹 & SNS상의 사과 가격과 관련된 비정형 데이터를 수집 및 분석하여, 앞으로의 가격 동향을 파악함으로써 사과 출하하시기 결정과 판매가 책정에 참고자료로 활용 가능

그림 29 ● 후지 가격 추세 그래프



자료 : 경상북도·미래창조과학부·한국정보화진흥원, 2013

02

기상 빅데이터 분석을 통한 효율적인 위험기상 예측

- 기상청

가. 추진 배경

❶ 일기예보의 현황 및 문제점

- 현재 기상예측에 활용되는 데이터는 관측데이터, 현재와 유사한 과거 데이터, 수치모델 활용 데이터에 한정
- 각 지역의 구체적인 기상예측을 위해, 방대한 양의 기상 데이터를 전문가의 지식과 추론에 의존
- 방대한 데이터를 인간 해석에 의존하기에는 한계점이 있음 예) ECMWF²⁾ 1.125도 자료 → 특정시각의 기상자료가 약 9천만($320 \times 161 \times 254 \times 7$)개 이상, 320×161 그리드, 254개의 속성, 9개 등 압면
- 위험 기상 예측의 경우, 유사한 기상상황에서도 지역적으로 매우 다른 특성을 보이며, 각 지역에 따라 어떠한 위험기상이 발생할지 모두 예측하기는 어려움

❷ 위험기상 예측

- Support Vector Machine(SVM)³⁾은, 비호우 상태의 일기도를 학습하여, 새로운 데이터가 들어왔을 때 그 일기도가 호우인지 비호우인지 분류

2) European Centre for Medium-Range Weather Forecasts : 유럽중기예보센터

3) Support Vector Machine(SVM) : 데이터를 분리하는 초평면 중에서 데이터들과 거리가 가장 먼 초평면을 선택하여 분리하는 방법

- SVM의 결과에 근거하여 위험, 비위험 기상인지를 구분하여 출력
- 기상 예측의 정확성을 높이기 위한 모델을 구성

나. 추진 내용

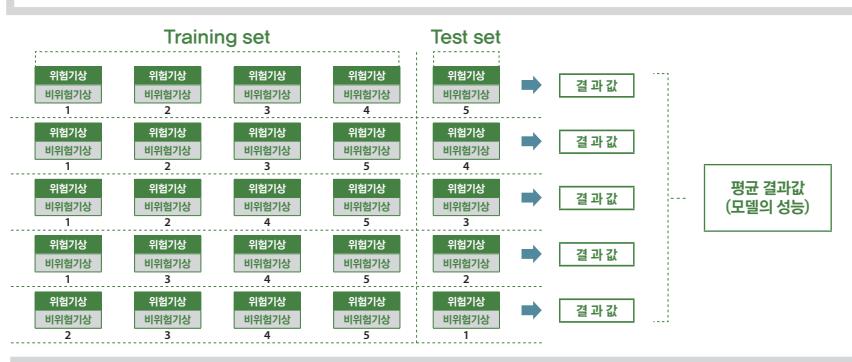
❶ 연구모형 구성

- 과거 기상데이터를 기반으로 7개의 위험기상 호우, 풍랑, 대설, 강풍, 한파, 폭염, 낙뢰의 발생 가능성을 예측하는 각 지역별 맞춤 모델 생성⁴⁾
- 예측 모델 생성 : 위험기상 별 권역을 남한지역을 4~16권역으로 나눈 후, 효율적인 속성을 선택하여 6시간, 12시간, 24시간 후를 예측 → 총 198개의 모델 생성

❷ Support Vector Machine

- 과적합(Overfitting) 문제 조정 가능
- 주어진 조건에서 최적의 글로벌 해를 구함
- 많은 특징값(Feature)을 사용하여 모델 생성 가능

그림 30 • Support Vector Machine

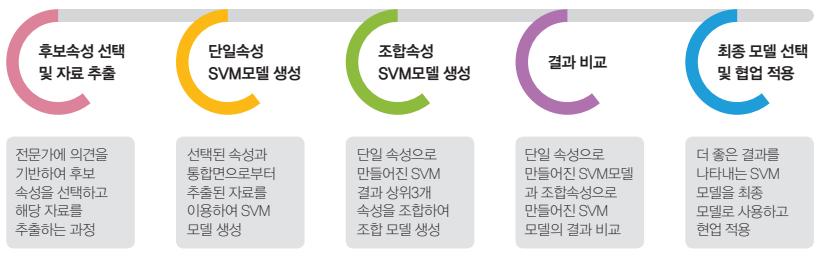


자료 : 이재동, 2012

4) 수행 기관 : 성균관 대학교 정보 및 지능 시스템 연구실

그림 30 • 제안 모델

Overview



자료 : 이재동, 2012

▣ 창조적 항공기상서비스 제공

- 기상청, 국토교통부, 국적항공사, 인천국제공항, 안전행정부 등 5개 기관의 정보 공유와 협조를 통해 항공사가 필요로 하는 수요자 맞춤형 기상정보 서비스 제공 가능
- 항공기 운항 종사자가 항공편명을 입력 시, 해당 항공기의 이륙, 운항, 착륙 등에 필요한 기상정보 (기상예보, 공항실황, 낙뢰, 화산재, 태풍, 비행고도별 기온과 바람, 극지방 운항 시 우주기상) 등을 무료로 확인

다. 성과 및 시사점

- ▣ 단일속성모델 분석을 통한 조합속성 모델을 생성하고, 생성된 모델 중 최적의 모델을 선택
- ▣ 대부분의 위험기상에 대하여 65~80%의 정확성을 보여 위험기상을 알리기에 적합
- ▣ 이전까지는 국내 항공사가 운항에 필수인 전 세계의 기상 정보를 외국 기상업자를 통해 구매했으나 (사용료 : 한 항공사 당 연 5천만원 추정), 해당 정보를 국내 항공사에 무료 제공함으로서 항공사 예산을 절감하게 하고, 일반 국민까지 활용 가능하게 확대 개발 예정

03

안전한 금융서비스를 위한 여신상시감시시스템

– 금융감독원

가. 추진 배경

- 저축은행의 과잉 실적 중심 운영에 따른 불법 및 부실 여신⁵⁾ 사례가 증가해 많은 서민이 피해를 보는 사태가 발생
- 금융감독 혁신방안 이행 과제의 일환으로 저축은행의 여신관련 통합 정보를 이용하여 불법 및 부실 혐의 여신을 사전에 탐지하고 사전에 방지

나. 추진 내용

- 전체 저축은행이 취급한 대출내역 및 대주주 정보, 신용평가사의 기업신용정보 등을 매월 입수하여 이를 분석, 16개 유형으로 구분하여 데이터베이스로 만듦

5) 여신 : 금융기관에서 고객에게 돈을 빌려 주는 일

그림 32 • ALESS⁷⁾ 업무처리 흐름도

자료 : 금융감독원, 2012

표3 이상 징후 여신(협의여신) 점검 항목

구분	점검항목
불법	대주주 등에 대한 여신, 개별차주 ⁷⁾ 신용공여 한도초과, 동일차주 신용공여 한도초과, 거액신용공여 한도초과 등 10개
부실	자산전성 부당분류, 프로젝트 파이낸스(PF) 대출의 일반대출 부당분류 등 2개
기타	대환 ⁸⁾ 협의, 증액대출협의 등 4개

자료 : 금융감독원, 2013

❷ 숨어있는 불법협의 여신 탐색 가능

- 저축은행의 채무관련인 간 보증 · 담보내역, 신용평가사의 계열관계 등 관계회사 정보 등을 활용한 연관성 분석을 통해 대주주 신용공여 등 은폐된 불법 협의여신을 찾아냄

6) 여신감시시스템(ALESS : Advanced Loan Examination Support System) : 매월 모든 저축은행으로부터 140만 계좌에 달하는 전체 여신현황자료를 제출받아 상시적으로 불법 · 부실 이상징후여신을 추출하는 고도화된 혼미경식 감시 도구(금융감독원, 2013)

7) 차주 : 돈을 빌려 쓴 이

8) 대환 : 신규 대출로 기는 대출을 갚는 행위

다. 성과 및 시사점

- ❷ 금융감독원의 현장 검사를 체계적으로 지원하여 보다 집중적이고 효과적인 여신 검사가 가능
 - 여신상시감시시스템을 활용, 상시감시와 현장검사를 유기적으로 연계하여 선택과 집중을 통해 검사 업무의 효율성을 제고
- ❸ 저축은행의 준법의식 제고로 스스로 불법 및 부실 여신 취급을 자제하도록 하는 예방효과
 - 여신상시감시시스템을 통해 추출된 건전성 부당 분류 혐의 여신의 규모가 2012년 12월 1조 6,019억원에서 2013년 6월 7,135억원으로 약 55% 감소(금융감독원, 2013)
- ❹ 여신상시감시시스템을 활용하여 365일 상시감시
 - 저축은행에 대한 정기 현장검사는 점차 축소하고, 상시 감시 과정에서 불법 · 부실 혐의가 드러나면 즉시 집중 검사를 실시하여 신속하고 정확한 대응 가능

04

특수질병 및 질환 예방 및 관리

- 보건복지부(국립보건연구원)

가. 추진 배경

❶ 보건 의료 산업의 핵심 재료로 인체 자원이 급부상

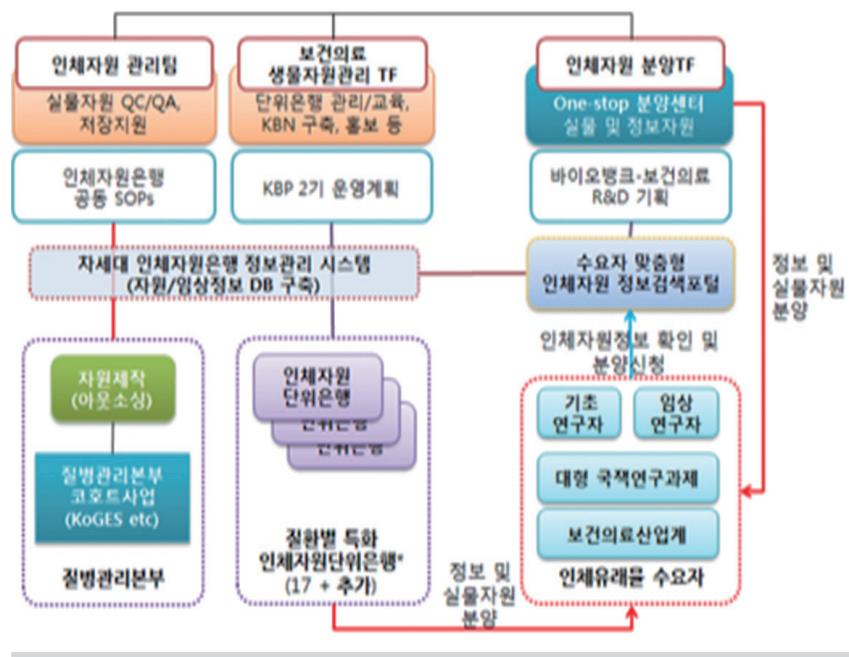
- 질병 지표 발굴 및 조기 진단을 위해 대규모 인체 자원이 활용됨에 따라 질병 예방 및 맞춤 치료 등 삶의 질 향상에 기여하기 위해 인체 자원 정보를 공유

나. 추진 내용

❷ 수요자 맞춤형 분양시스템 구축

- 1기 사업에서 수집한 50 만 명 분의 병원별로 특성화된 질환 집단에 속하는 인체 자원을 관리, 이를 활용할 연구자에게 적극 분양하기 위해 One-Stop 분양이 가능한 온라인 분양 포털을 마련
- 2013년에 2기 사업으로 진행됨

그림 33 • One-Stop 분양 포털 추진 체계



자료 : 보건복지부, 2013

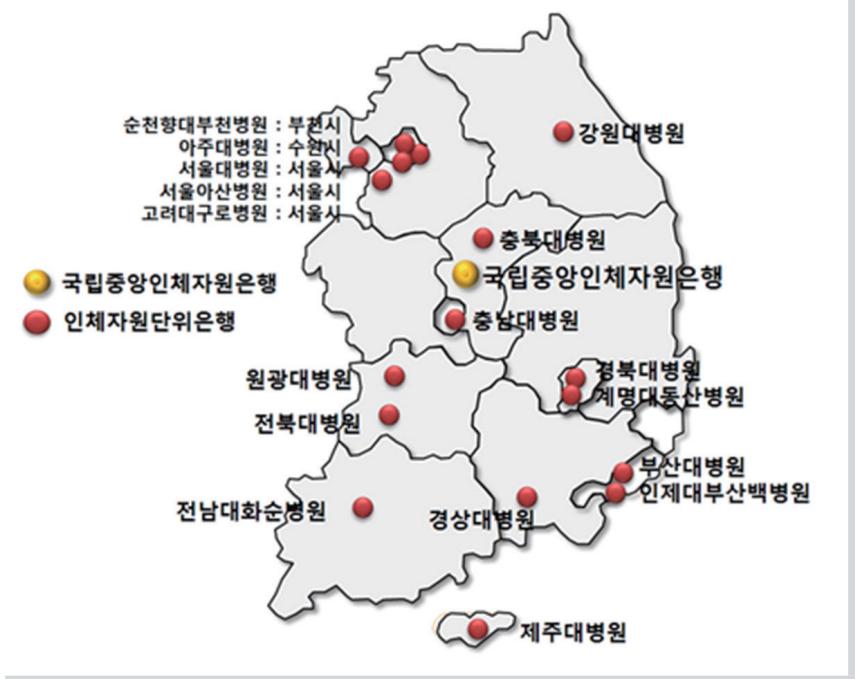
 관계 부처별 유전체 지원 보유, 활용현황을 '국가 생명연구지원 정보센터'(KOBIC)에서 중점관리

- 2021년까지 10만 명의 한국인 유전체 자원을 확보, 인체자원연구지원센터를 통해 맞춤의료, 예방치료 등 보건의료 신 성장 분야 지원

인체자원의 품질 표준화를 위한 운영 프로토콜 기반 관리

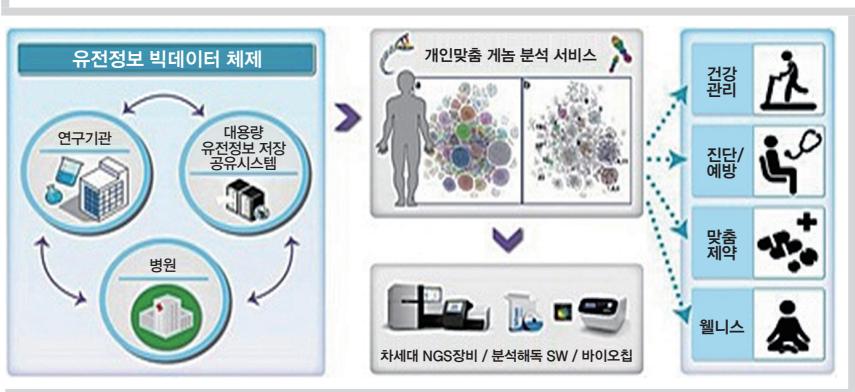
- 단위은행은 인체 자원의 품질 표준화를 위해 표준화된 은행 운영 프로토콜(규약)을 기반으로 인체 자원을 수집, 보관 및 분양
 - 수집된 자원을 기반으로 다양한 질병연구, 바이오마커 개발, 정도관리지표 물질 개발 등 인체 자원 활용화 연구 수행

그림 34 • 인체자원은행 전국 현황(1개 중앙은행, 17개 단위은행)



자료 : 보건복지부, 2013

그림 35 • 유전정보 빅데이터 시스템 흐름도



자료 : 쿠기뉴스, 2013

다. 성과 및 시사점

국가 주도의 보건 의료 분야 생명 연구 지원의 글로벌 경쟁력 기반 확보

- 생명 연구 지원의 체계적 수집과 정보 표준화, 통합 네트워크 구축 가능
- 질병관리본부, 단위은행 뿐만 아니라 민간 인체자원은행의 보유자원까지 표준화 하여 보관, 활용하는 표준화 관리 정책 개발 도입 가능

미래 바이오 산업의 신 성장동력으로 자리매김 가능

- 연구자가 요구하는 자원을 분양받을 수 있도록 파생 시료를 공급하고, 고부가가치 제품 개발에 필요한 인체 자원 확보와 연계 확대

05

통합형 재난안전 정보 네트워크, 스마트 빅 보드 (첨단 재난 상황실)

- 안전행정부(재난안전연구원)

가. 추진 배경

❶ 재난 상황에 대한 통합 데이터가 부족

- 그간 재난·안전 사고 발생 시 주로 문서나 유선을 통한 상황 보고에 그친 현장 정보와 언론사나 지자체에서 제공된 영상 정보를 수동적으로 이용함
- 재난상황을 실시간으로 파악하고 능동적으로 분석하여 국민들의 신속한 안전 확보에 도움을 주기 위한 체계적인 서비스에 대한 필요가 꾸준히 제기됨

나. 추진 내용

❷ 스마트 빅 보드⁹⁾를 통해 기상, 재난이력 및 국내외 재난정보의 정보 공유를 통해 재난현장 상황 정보 취득 가능

- CCTV·무인항공기·스마트폰·인공위성영상 등 다양한 첨단 기술을 통해 지상은 물론 항공에서 제공되는 생생하고 입체적인 재난현장 상황 조회 가능
- 종합적인 원스톱 재난상황 모니터링 : 위치기반 SNS 정보 및 트위터 등 재난현장에서 제공되는 다양한 정보를 이용하여 신속한 대응이 가능
- 빅데이터 정보 : 재난관리에 활용성이 높은 실시간 트윗 정보, 과거 재난이력 및 원인분석결과, 위성영상, 시뮬레이션, 관련 웹사이트 등을 빅데이터 메뉴로 분류

9) 스마트 빅 보드 : 재난 및 안전사고 시 현장 중심의 모든 가능한 정보 네트워크를 가동하여 위험상황을 분석하는 최첨단 스마트 재난 상황실

하여 예방, 대비, 대응, 복구의 재난 관리의 전체 프로세스에 과학적이면서도 직관적인 재난관리 수행 지원

▣ 상황실 개선 방향

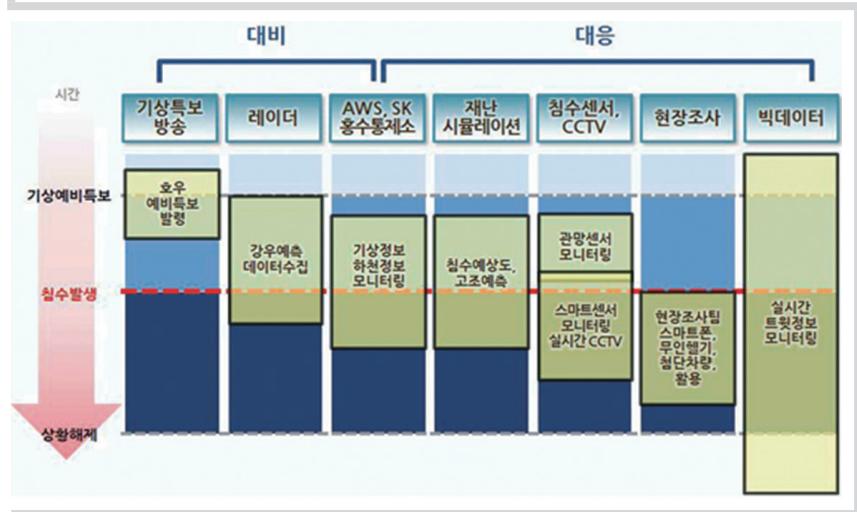
- 기존 재난관리 관련 기관에서 개별적으로 관리·서비스하고 있는 정보를 연계하여 하나의 프레임 안에서 위치 기반 서비스 (LBS : Location Based Service) 형태로 표출하여 다양한 유형의 재난에 신속하게 대응

표4 상황실 개선 방향

구분	현재	개선
정보형태	기상, CCTV 등	지능형 CCTV, SNS, 재난지도 등
정보분석	경험적 위험성 판단	시뮬레이션 등 과학적 분석
위험경보	기상청 등 단순정보 의존	지역특성을 고려한 스마트 재난경보
현장경보	TV, 뉴스 등 언론정보	위성, UAV, MMS(이동차량) 등
정보통합	각 부처별 재난정보 관리	부처간 정보통합 구축
정보표출	보고서 방식의 단순 정보표출	위치기반 전자지도 표출
민관협력	정부중심의 일방적 대응	SNS 등 국민 ↔ 정부간 쌍방향 대응

자료 : 안전행정부, 2013

그림 36 • 스마트 빅 보드 운영 시나리오



자료 : 최우정 외, 2013

그림 37 • 스마트 빅 보드 메인화면



자료 : 최우정 외, 2013

다. 성과 및 시사점

❶ 다양한 유형의 재난에 신속 정확하게 대응

- 상황 판단에 필요한 정보를 신속하게 제공하고 최적의 상황대응 의사결정을 하는데 도움을 줌

예) 2013년 4월 산대저수지 붕괴사고와 7월 아시아나 항공기 사고 시 언론보다 먼저 위험상황 전파가 이루어졌던 사례는 스마트폰 3천 5백만 대 시대의 미래형 재난관리의 청사진을 제시(국립재난안전연구원, 2013)

❷ 국민생활안전지도의 제작·보급에 도움

- 재난 상황을 전자 지도로 표출하고, 재난 예측 및 피해 시뮬레이션 결과를 바탕으로 과학적 재난 의사 결정 지원 가능

06

맞춤형 정보 유통 서비스

– 안전행정부(행정정보공동이용센터)

가. 추진 배경

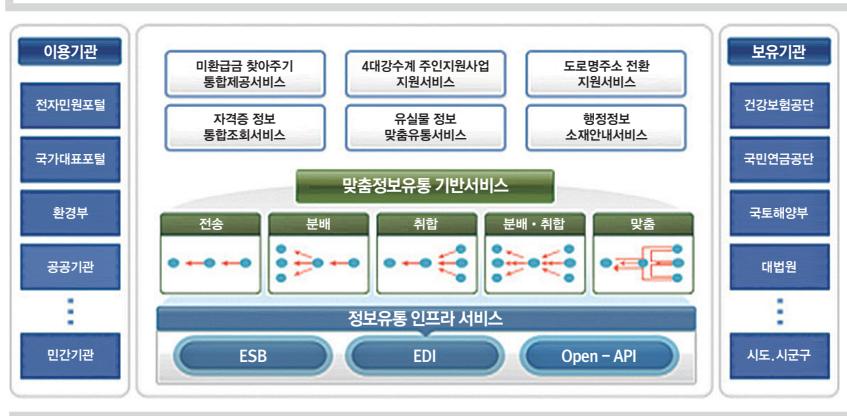
- ❶ 민원 관련 정보를 각 기관이 개별적으로 관리하여 발생하는 다양한 문제들에 대한 통합적인 해결책 개발 필요
 - 국민의 불편 최소화, 개인정보 보호 및 예산 절감 필요
 - 국민들이 각종 민원 신청 시에 여러 개의 각종 구비 서류를 지참하고 관공서를 방문하여 복잡한 절차를 통해야 하는 불편 : 2008년 기준 5,037종의 민원사무와 관련된 구비 서류가 3,680여종(행정정보공동이용센터, 2013)
 - 기관별 정보의 중복 수집 및 관리로 낭비되는 예산 감축 필요 : 연간 4억 4천만통의 서류가 각 행정기관에서 발급되어 공공·금융기관 등에 제출되고 있어 예산 낭비, 보관 공간과 행정력의 낭비가 발생(행정정보공동이용센터, 2013)
 - 다량의 개인정보를 각 기관이 개별 관리하여 개인정보에 대한 보안 문제 발생 우려 (행정정보공동이용센터, 2013)

나. 추진 내용

- ❷ 맞춤 정보 유통 기반 서비스¹⁰⁾ 개요

10) 행정정보공동이용시스템의 고도화된 정보유통 인프라를 통해 다수 보유기관의 정보를 융합하여 생활공감형 통합 정보를 일괄 제공하는 서비스

그림 38 • 맞춤 정보 유통 기반 서비스 개요



자료 : 행정정보공동이용센터, 2013

● 미환급금 찾아주기 통합서비스

- 국세청, 안전행정부, 대법원, 국민건강보험공단, 국민연금공단, 근로복지공단 등 기관별로 따로 관리되고 있는 미환급금 시스템을 정보맞춤서비스를 통해 연계하여 국민들이 미환급금 유무를 한 곳에서 실시간으로 확인하고 환급신청 가능

그림 39 • 10종 미환급금

번호	연도	미환급금 정보명	보유기관
1	2010 (4종)	국세미환급금	국세청
2		지방세환급금	안전행정부, 서울시
3		보관금	대법원
4		송달료	대법원
5	2011 (6종)	건강보험료 과오납금	국민건강보험공단
6		본인부담환급금	국민건강보험공단
7		본인부담상환액 초과금	국민건강보험공단
8		기타징수금 과오납환급금	국민건강보험공단
9		국민보험료 과오납금	국민연금공단
10		고용산재보험료 과오납금	근로복지공단

자료 : 행정정보공동이용센터, 2013

그림 40 • 유실물 정보 맞춤 정보유통서비스



자료 : 행정정보공동이용센터, 2013

❷ 유실물정보 맞춤 정보 유통 서비스

- 한국철도공사, 서울 지하철 9호선에서 관리하고 있는 분실물, 습득물 정보를 경찰청 유실물종합관리시스템과 연계하여 국민들이 분실한 유실물을 신속하고 편리하게 찾을 수 있도록 지원

❸ 4대강 유역 주민생활서비스

- 4대강 유역 주민지원을 위한 선정업무를 수행하기 위해 국토교통부 토지 이용계획 확인서, 대법원 부동산 등기전산정보, 안전행정부 주민등록전산정보를 맞춤, 유통하여 환경부에 전달하는 서비스

다. 성과 및 시사점

- ❶ 민원 관련 개인정보의 통합 관리로 예산 절감 및 보안 문제 해결
- ❷ 국민들의 민원 신청에 필요한 서류 구비에 따른 불편 최소화

07

지리정보 통합한 지리적 프로파일링 시스템 구축 (GeoPros)

- 경찰청

가. 추진 배경

❶ 지리적 프로파일링¹¹⁾ 시스템¹²⁾ 개발의 필요성 대두

- 방대한 범죄수사 자료가 축적되고 수사기법이 발전되어 왔으나, 한국 실정에 맞는 지형공식이 없어 범죄수사의 범위를 줄여줄 지리적 프로파일링 시스템 개발의 필요성이 생김

나. 추진내용

❷ 범죄 유형에 따른 범죄 위험 지역을 한 눈에 알아보게끔 시각화

- 각종 범죄의 유형에 따라 지역 및 시간대별로 범죄다발지역과 위험도를 지도에 각기 다른 색으로 표시

❸ 추가 범죄 가능성을 예측하여 우선 수사 및 예방

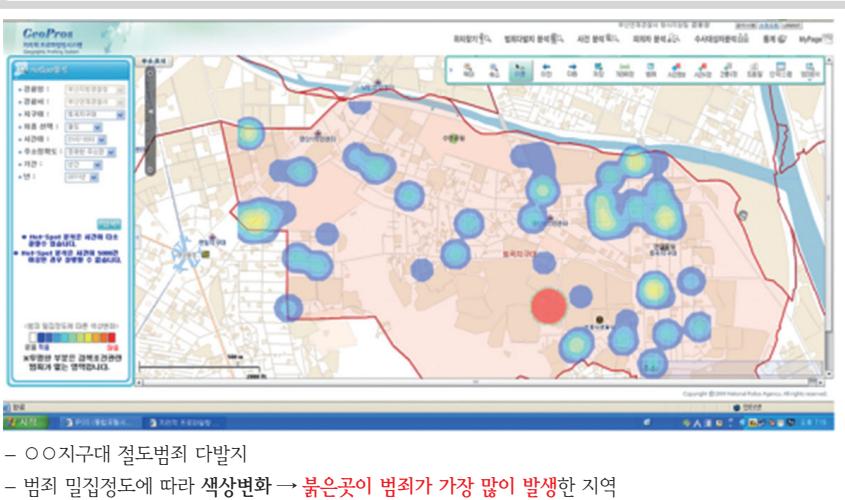
- 기존 발생한 범죄의 데이터를 통합 수집 및 분석하여 추가 범죄의 가능성과 방향성을 찾고 우선 수사 대상자 및 대상 지역 선별

11) 지리적 프로파일링 : 범죄현장에서 수집된 데이터들을 이용하여 범인의 행동 혹은 심리적 특성을 추론하여 용의자의 확률이 높은 인구통계적 특성, 단서를 도출하는 것(신상화, 2009, p.129)

12) 지리적 프로파일링 시스템 : “다양한 공간통계분석기법을 경찰의 범죄수사데이터에 적용, 범죄위험지역 예측을 통한 방법 전략 수립 및 연쇄 범죄자 거주지 예측을 통한 수사 활동 전개가 가능한 시스템”(경찰청, 2013)

- 연쇄 범죄의 발생 가능성이 높은 장소와 시간을 예측하고 용의자가 거주할 만한 지역을 예측
- 범죄위험도에 따라 순찰차를 배치하고 순찰 경로를 선정
- 성폭력 등 범죄에 대한 지역별, 시기별 맞춤형 치안대책 수립

그림 41 • 지리적 프로파일링 시스템 활용 범죄 다발지 분석 예시



자료 : 부산연예경찰서, 2012

그림 42 • 범죄예측 지리적 프로파일링 시스템 활용



자료 : 경찰청 블로그

다. 성과 및 시사점

❷ 강력 범죄 발생의 감소 및 예방

- 지리적 프로파일링 시스템을 통해 분석한 자료를 토대로 2012년 8월 1개월간 범죄 다발지(5대 범죄) 중심 형사기동차량순찰 및 필요지점에 대한 거점근무를 실시한 결과, 해당 범죄가 전년 동기간보다 10.5% 감소(전년 동기간 296건, 2012년 8월 265건) (부산연제경찰서, 2012)

❸ 순찰 현장의 대응력 강화로 주민 안전 고취

- 향후 112 신고출동시스템 및 순찰차 내비게이션과 연계하여 시간별, 지역별 범죄 위험지역을 순찰차에서 바로 확인하여 전략적 순찰활동 전개 가능(안전행정부, 2013)

❹ 주민의 관심도를 높이고 주민 스스로의 범죄 예방 유도

- 범죄예방지도 등 생활치안 정보를 생산하여 국민에게 공개함으로써, 지역 치안에 대한 관심을 높이고 범죄를 예방하는 효과 기대(안전행정부, 2013)

08

국가공간정보 통합데이터베이스 구축

– 국토교통부

가. 추진 배경

- ❶ 유비쿼터스, 모바일, 가상공간 등 정보의 융·복합 기술 발전에 따른 공간정보 수요증가
- ❷ 산재해 있던 다양한 분야의 GIS (지리 정보 체계) 정보를 연계 통합 필요성 제기
 - 중앙부처, 지방자치단체, 공공기관 등이 토지, 도로, 지하시설물, 통계 등의 정보를 각각 구축, 운영함에 따라 상호 호환성이 떨어지고, 중복으로 개발하여 예산이 낭비되고 시스템 간 연계체계의 미비로 활용도가 떨어짐
 - 다수의 기관에서 개별적으로 구축, 활용되던 공간정보를 연계, 통합하여 행정업무에서 공동으로 활용할 수 있게끔 공간정보통합체계를 구축

나. 추진내용

- ❸ 중앙부처 공간정보시스템 연계 지속 추진
 - 국가공간정보 공동 활용의 통합 관리 기관 역할을 강화하기 위해서 25개 기관 76개 시스템 연계 완료
 - 246개(광역 17, 기초 229) 지자체 확산을 통해 국가공간정보 통합체계 정착 완료 (국토교통부, 2013)

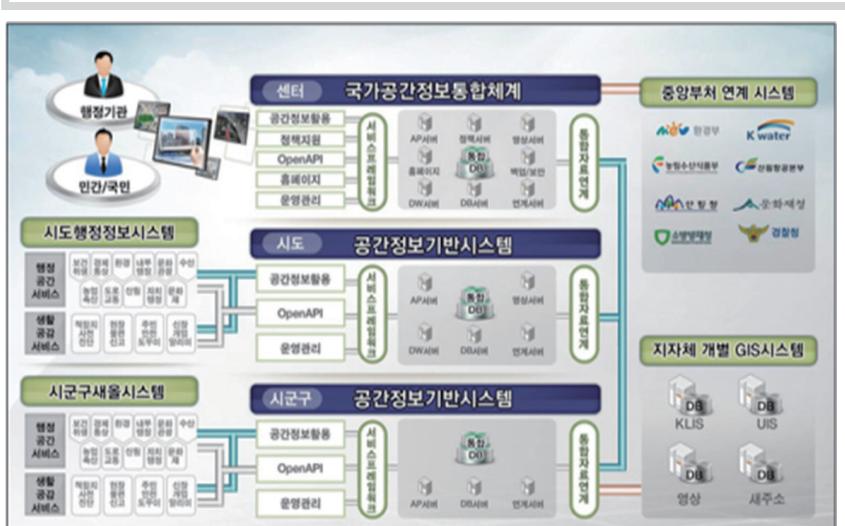
▣ 민간 활용수요가 높은 공간정보의 공개

- 국가공간정보센터를 통해 행정정보, 수자원/해양, 환경, 산림, 보전지역, 토지 정보, 지형, 지질, 관광/문화 등 9개 분야 31종의 정보 공개에 이어 2013년 7월 택지정보, 도시계획정보, 등산로정보, 사업지구정보, 국가지명, 해안선정보, 국가지명, 교통CCTV, 국가교통정보 등 16종의 공간정보 추가 개방(국토교통부, 2013)
- 포털, 통신사, 내비게이션 회사 등이 영업점 설치, 부동산개발, 길안내 및 지도 서비스 간신이나 최신 정보 구축에 활용 가능(국토교통부, 2013)

▣ 공간 빅데이터의 융합된 데이터 베이스를 바탕으로 새로운 가치 창출 가능한 분석모형 개발

- 3차원 분석, 시계열 분석, 기반시설 입지분석 등
- 부동산, 교통, 안전, 복지 등 구체적인 행정 분야에 공간 빅데이터 플랫폼을 적용하여 다양한 서비스 개발 및 제공(국토교통부, 2013)

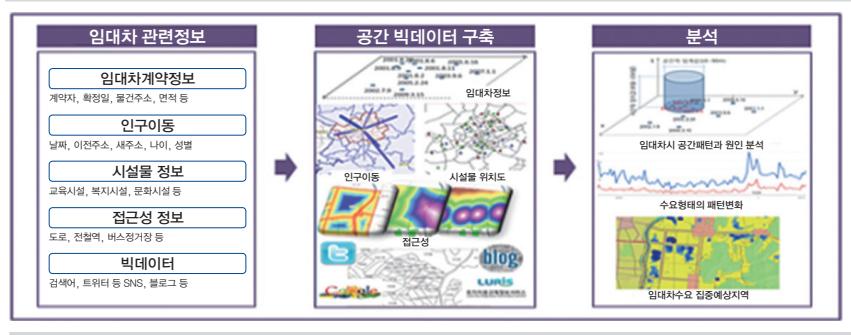
그림 43 • 국가공간정보 통합데이터베이스 구축





자료 : 국토교통부 홈페이지

그림 44 • 공간 빅데이터를 활용한 ‘부동산 임차시장 수요패턴 분석 및 대응방안’



자료 : 국토교통부 홈페이지

다. 성과 및 시사점

❶ 분산 관리 되던 국가공간정보를 통합하여 필요한 국민에게 제공

- 공공기관의 행정 · 공간정보와 민간 정보 간 통합을 통한 생활 공감형 서비스 마련

❷ 보다 정확한 국토계획수립이 가능하게끔 지원

- 국가공간정보 플랫폼에 공간정보와 행정정보를 융합하여, 주택 수급이나 교통 등 인프라 구축 등의 행정서비스의 효율성과 질 향상 가능

09

고용정보통합분석시스템 구축

- 한국고용정보원

가. 추진 배경

❶ 고용정책을 위한 현황정보의 적시 제공 필요

- 과거 고용 관련 이슈 분석에 자체되는 시간 때문에 적시에 효과적인 연계 분석 어려움
- 개별 운영되는 시스템에 자료를 별도로 요청할 때 소요되는 시간이 지나치게 길어 시기적절한 분석 결과 활용이 어려웠음

❷ 데이터의 품질을 높여 대국민 서비스 개선 필요

나. 추진 내용

❶ 고용정보 전산망 통합 관리

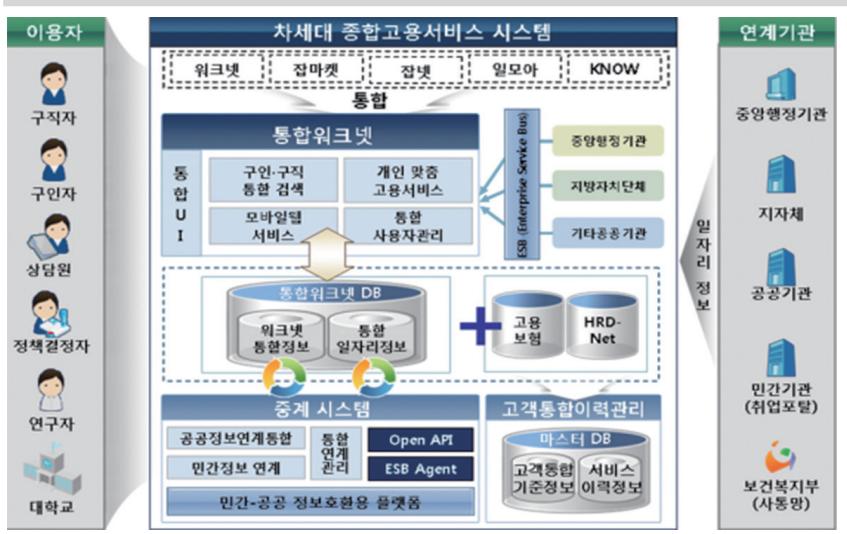
- 고용보험전산망, 워크넷, 직업훈련전산망, 외국인고용관리전산망 등 11개 전산망 보유 및 관리

❷ 신속하고 다양한 분석을 통해 고용정보의 적시 제공(한국고용정보원, 2013)

- 연계분석 : 고용보험, 취업알선(워크넷), 직업훈련 등의 다양한 고용정보를 연계 분석한 자료 파악
- 고용동향 : 워크넷, 외국인, 자활, 사회적 일자리, 구인업체개척사업, 경제활동 인구조사 등 관련 통계 자료 추출 및 분석

- 고용보험 : 사업장, 피보험자, 모성보호, 지원금, 실업급여 관련 통계 자료 활용
- 직업훈련 : 실업자, 재직자, 고용촉진, 지도점검, 훈련교사 관련 통계 자료 활용
- 부서현안자료 : 고용노동부 각 팀 별 해당 통계 자료 제공

그림 45 ● 차세대 종합 고용 서비스 시스템 구성도



자료 : 공개SW포털

다. 성과 및 시사점

● 민간의 노동시장 관련 정보 직접 추출 이용 가능

● 정부 고용정책의 실효성 증대

- 축적된 데이터베이스를 활용해 신뢰성 있는 정보를 제공, 다양한 연계분석으로 고용정책 실효성 높임

예) 고용정보시스템을 통해 개인 특성, 구직자 고용 변경 추이 및 현재 상태 파악부터 사업장 특성, 고용 유지 기간 등의 통합 정보 확인 가능

예) 직업훈련시스템을 통해 개인특성이나 사업장 특성 등의 특성별 정보 취득 : 취업자의 재직자 훈련 이수 여부, 실직자의 실업자훈련 프로그램 참여 여부 확인 및 사업자의 고용 유지 추이 파악 가능

10

역외탈세 정보공유

- 관세청, 국세청

가. 추진 배경

❶ 탈세 적발의 강화 가능성 발견

- 역외탈세¹³⁾와 불법외환거래는 동시에 발생하는 특징을 보임
- 해외발생 소득, 신고누락에 따른 역외탈세를 적발하는 국세청과 불법외환거래를 파악하는 관세청 간에 일시적, 소극적 정보공유가 아닌 지속적, 적극적 정보 공유를 통해 불법외환거래 및 역외탈세에 대한 단속 효과를 동시에 높이는 것이 가능

나. 추진 내용

❷ 두 기관 간 지속적, 적극적 과세 정보 공유

- 관세청과 국세청이 각자 보유 중인 신고자료 등의 과세 정보 교환을 통해 정보 협력 체제 강화
- 국세·관세 탈루 등 혐의 정보를 '외환 거래 감독기관 협의회'를 통해 정기적으로 교환
 - 예) 2013년 9월 현재 관세청은 불법외환거래 조사 중 발견한 역외탈세 등 내국세 탈루 혐의 정보를 국세청에 제공, 국세청은 관세탈루 및 채권미회수 등 불법외환거래 혐의 정보를 제공해 정밀 분석

13) 역외탈세 : 세금을 피하기 위해 해외의 조세피난처 등을 이용한 탈세행위(SBS CNBC, 2013)

다. 성과 및 시사점

행정 효율성과 조세형평 증대

- 인력이나 예산의 추가 투입 없이 정보 공유만으로 행정효율성과 조세의 형평을 증대시키는 정부 부처 간 대표 협업 사례로 떠오름
- 지능적 역외탈세에 효과적으로 대처

세수 증대 및 국고 손실 방지(안전행정부, 2013)

- 탈세 혐의정보 공유의 획기적 증가로 연간 약 900억원 세수증대
 - 과세 기초정보 공유의 확대로 연간 약 1,000억원 세수 증대
 - 복지부 등 다양한 기관과의 정보 공유 확대를 통해 약 485억 원 상당의 국고 손실 방지
- * 2013년도 <정부 3.0 우수사례 대상>으로 선정

해외 협력으로 국제 공조 네트워크 구축으로 발전

- 미국, 영국, 호주와 함께 공동 조사하는 역외탈세정보 공유 네트워크 구축

그림 46 • 역외탈세 적발 국제공조 네트워크 구축



자료 : 광주지방국세청, 2013

11

빅데이터 분석을 통한 2013 국민 인식 변화 분석

– 문화체육관광부

가. 추진 배경

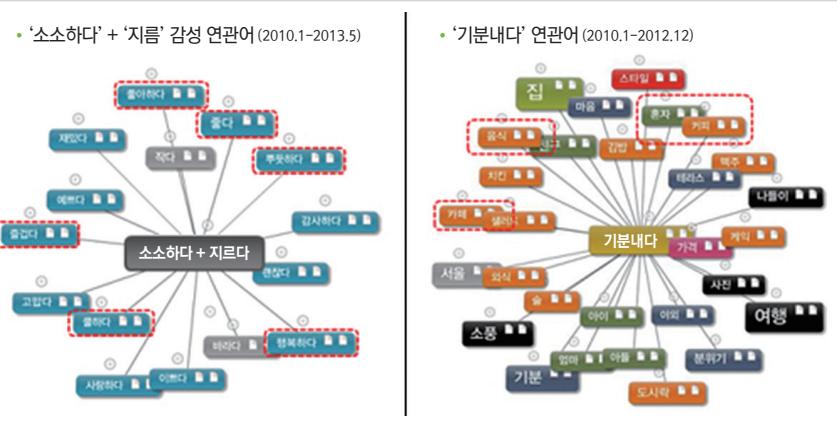
- ❶ 문화, 체육, 관광 관련 정책 수립에는 간접적인 수요와 연관된 국민들의 여가에 대한 인식 현황 파악이 중요
- ❷ 인터넷 및 스마트폰의 보급과 동시에 최근 국민들의 생활에 변화가 있었을 것으로 예측 가능하나 실제 데이터로 분석한 바가 없어 그 필요성 대두

나. 추진 내용

- ❶ 온라인 뉴스 및 SNS(소셜 네트워크 서비스) 등의 정보를 통해 국민의 라이프 스타일 조사
 - 다음소프트와 협력하여 2011년 1월 ~ 2013년 5월까지 29개월 간 트위터, 블로그, 온라인 뉴스 등을 통해 36억여 건의 메시지를 분석
- ❷ 해당 기간 중 언급이 증가한 키워드와 연관어 분석
 - 키워드 : 현재, 일상, 퇴근 후, 소소하다, 지르다, 혼자 등 소소한 일상을 소중하고 행복하게 여기고, 작은 물건들을 사며 뿌듯해하며 영화 감상, 여행, 커피 등을 혼자 즐기는 추세도 증가하는 것으로 나타남
 - 정책 관련 : 일자리, 사회복지, 무상교육, 영어 조기교육 등 일상생활에 밀접한 연관이 있는 분야에 관심

- 박근혜 정부 4대 국정기조 관련 : 경제부흥, 국민행복, 문화융성, 평화통일 기반 구축 중 가장 많이 언급된 키워드는 ‘국민행복’(48% 차지)(문화체육관광부, 2013)

그림 47 • 키워드 연관어 분석



자료 : 문화체육관광부 (2013)

다. 성과 및 시사점

- ▣ 국민이 정부에 기대하는 요소를 파악함으로써 그에 부응하는 정책 수립 지원 가능

- 국민안전에 대한 관심, 튼튼한 안보에 대한 언급이 많았던 것으로 나타나 정부가 생활안전, 안보 등 최소요건을 담보해주기를 기대하는 것으로 나타남
 - 창조경제를 통해 경제에 활력을 주고 일자리를 만들어주기를 기대하며, 일자리 및 부동산 등 삶에 직접 영향을 주는 정책에 대한 관심이 높았음
 - 건강 관련 멀거리 안전에 대한 요구, 문화 관련 취약 계층에 대한 문화 대상 확대 등도 관심이 높아 관련 정책의 우선순위 설정에 참고 가능

참고문헌

- 경상북도(2013), 경상북도 공공분야 빅데이터 활용방안 연구
- 경찰청(2013. 2.7), 경찰순찰! 범죄 위험지역 위주로 집중!! 효율성 강화.
- 경향신문(2013.9.30.), 경북, 빅데이터 활용한 맞춤형 사과재배 정보제공시스템 구축
- 공개SW포털(2013), 한국고용정보원 – 공개SW 기반으로 고용정보 시스템 서비스 개선
- 국세청(2013), 국세청, 관세청과 ‘정부3.0 우수사례’ 대상 수상 문화체육관광부(2013.6.27), 혼자, 일상에서 소소한 행복을 느끼다 – 36억여 건의 빅데이터 분석을 통해 본 2013 국민 인식 변화
- 국토교통부(2013), 국가공간정보통합체계 국토교통부(2013.6.30), 국토부, 민간활용도 높은 국가공간정보 본격개발 시작했다
- 금융감독원(2012), 저축은행 불법 · 부실혐의 여신 주출을 위한 상시감시시스템 구축
- 금융감독원(2013), 저축은행 여신상시감시시스템 가동을 통한 불법행위 적발 및 조치
- 더아이엠씨(2013.12.18.), ‘국내 빅데이터 활용사례’
- 대한측량협회(2013), 공간정보 빅데이터체계 구축, Special 2, Surveying & Mapping Magazine.
- 데일리그리드(2012), 한국고용정보원 고용정보통합분석시스템 구축사례.
- 데일리그리드 리서치센터 소프트웨어그룹.
- 매일신문(2012.12.24.), ‘근혜 사랑’ 영주사과…농부, 새누리 당사 전달
- 메디컬월드뉴스(2013.8.5), 제2기 한국인체자원은행사업 추진계획 확정
- 보건복지부(2013), 제 2차 한국인체자원은행사업 추진계획 – 2013~2015
- 부산연제경찰서(2012. 9. 6), 지리적 프로파일링 시스템 활용한 범죄다발지(hot-spot) 형사활동
- 세계일보 (2013.9.29.), 경북도, 빅데이터 활용 사과 작황정부 서비스 개발 나서
- 신상화(2009), 연쇄 강간범에 대한 지리적 프로파일링에 관한 연구. 경찰학논총. 4(2), pp. 127–160.
- 안전행정부(2013.4), ‘안전한 사회, 유능한 정부, 성숙한 자치로 행복한 대한민국을 만들겠습니다!’. 2013년 업무보고
- 안전행정부(2013.4.15), 모든 재난 정보, 한눈에 확인하고 즉시 대응
- 안전행정부(2013. 12), 관세청-국세청, ‘정부 3.0 우수사례’ 대상 공동 수상. 보도자료.

영남일보(2012.8.18.), 국내 1위 경북사과, 세계벽도 허물어라

이재동(2012). 기상 빅데이터 분석을 통한 효율적인 위험기상 예측 방법

장병열, 김영돈(2013), 빅데이터 기반 융합 서비스 창출 주요 정책 및 시사점. 정책초점. 과학기술정책. 제 23권 제 3호

전체남(2013), 농업 빅데이터를 활용한 병해충 발생 예측 : 영주사과를 대상으로. 한국데이터 사이언스학회 학술대회 및 정기총회

정부 3.0 블로그(2013). 정부3.0 창업과 일자리 창출 기대되는 이유 안전행정부 박찬우 제1차관에게 들어보니.

최우정, 이종국, 여운광(2013), 국립재난안전연구원 재난안전 “스마트 빅 보드” 개발 착수 – 재난관리 3.0을 선도하는 새로운 플랫폼. 물과 미래. 46 (9) : 73–78.

쿠키뉴스(2013.12.3), 개인유전정보로 맞춤형 의료서비스…정부, 유전정보 빅데이터 체제 구축 추진 한국 고용정보원(2013), 고용정보통합분석시스템.

행정정보공동이용센터(2013), 맞춤정보유통서비스 SBS CNBC(2013), [친절한 김기자] ‘뜨거운 감자’ 역외탈세란?

IV

빅데이터 활용 사례 Part 3.

민간사례 13종

- 
- 01** 빅데이터 분석을 통한 베노플러스겔 마케팅 (유유제약)
 - 02** SSD 온라인 버즈 마케팅 (삼성전자)
 - 03** C.S.I 프로젝트 (신세계몰)
 - 04** 고객 맞춤형 서비스를 위한 빅데이터 센터 (신한카드)
 - 05** 소비자 맞춤형 수요 매칭 (FNC 코오롱)
 - 06** 고객 맞춤형 매장 배치 (현대백화점)
 - 07** 맞춤형 고객 상품 추천 서비스 (GS샵)
 - 08** 맞춤형 여행 컨설팅 서비스 (롯데관광)
 - 09** 자동차 고객관계 관리 프로젝트 (현대자동차)
 - 10** 스마트 미디어 솔루션 서비스 (삼성전자)
 - 11** 모바일 빅데이터 맞춤형 서비스 (NHN)
 - 12** 스마트 통합물류 시스템 (CJ 대한통운)
 - 13** 언더라이팅의 예측모델 활용 방안 (푸르덴셜생명보험)

01

빅데이터 분석을 통한 베노플러스겔 마케팅

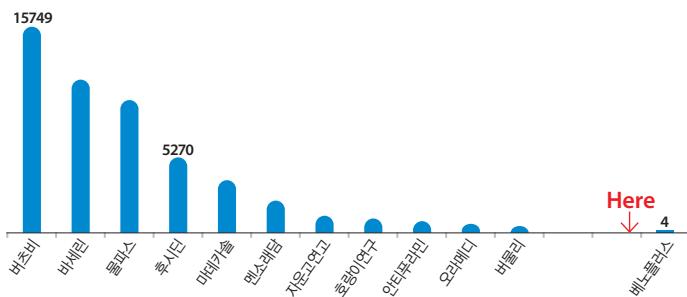
- (주)유유제약

가. 추진 배경

❶ 타회사와 비교하여 유유제약 제품의 인지도 개선 필요

- 과거 약사나 의사에 의존한 연고선택에서 현재에는 소비자의 주체적인 정보검색과 다양한 요구에 따른 선택이 증가
- 베노플러스겔¹⁴⁾의 다양한 효과성(멍, 봇기, 벌레 물린 곳, 생약성분 등)에도 불구하고, 최근 3년간의 데이터에 따르면 여전히 인지도는 미비함

그림 48 • 약품 인지도 그래프



자료 : (주)유유제약, 2012

- 빅데이터 활용을 통한, 유유제약의 베노플러스겔의 효능 관련 단어들이 소비자들에게 무엇을 중심으로 인식되어 있는지 분석이 필요

14) 베노플러스겔 : 유유제약의 명든데, 부은데, 타박상, 벌레물린데에 사용하는 무색, 무취의 겔 형태의 연고

나. 추진 내용

‘멍, 봇기’와 같은 주요 단어 대한 인식 분석

- 베노플러스겔의 주요 효능으로 멍과 봇기라는 단어가 소비자에게 어떤 식으로 각인되어 있어, 무엇을 떠올리는 가의 분석에서 출발
- ‘봇기’는 다이어트, 미용관련 봇기 제거를 중심으로 미용전문 한의원, 클리닉, 병원, 마사지와 강하게 링크가 되어 있음
- ‘멍’의 경우, 특정 연고와 강한 연관관계를 가지고 있지 않을 뿐만 아니라, 민간 요법과 더욱 가까이 링크

그림 49 ● ‘멍’ 들었을 때의 해결책들 그라프



자료 : (주)유유제약, 2012

● 마케팅 추진 내용

1) 민간요법과의 전쟁

- 제품에 있어서는 맨소래담과 버즈비가 경쟁 상대이지만, 소비자의 인식에서는 계란과 소고기 같은 민간요법을 극복해야 함
- 기존의 민간요법의 강한 인식을 멍 = 베노플러스로 전환시키는데 목적을 둠
- SNS 업체와 협력하여 ‘계란을 둘리는 것은 팔이 아프다’ ‘소고기는 비싸고 비위생적이다’라는 마케팅 메시지를 전파하고 ‘계란은 드세요. 멍은 베노플러스가 뺄게요’라는 광고 메시지를 개발

2) 마케팅 대상의 전환

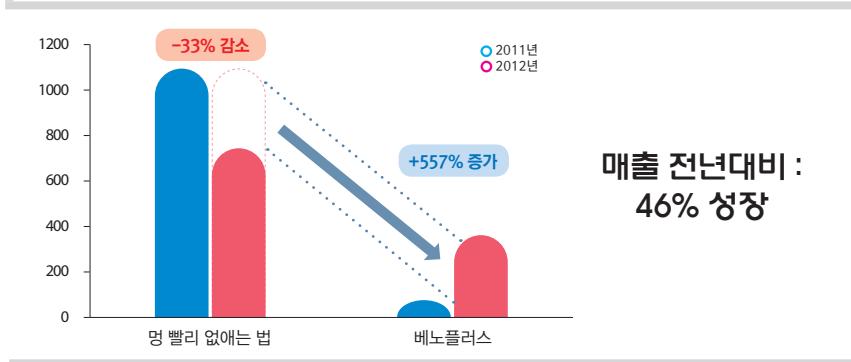
- 최근까지의 마케팅은 단지 어린이를 대상으로 이루어졌음
- 2008년에서 2012년까지 5년 간 지속된 소비자 트랜드 분석에 있어서 ‘멍’이라는 단어는 ‘여성’, ‘아이’, ‘환자’와 결부되어 나타남
- 그 결과 마케팅 대상이 ‘아이’와 ‘여성’으로 확장
- 그 밖에 마케팅 수단으로써 소아과에 대한 홍보와 기존 잡지에의 홍보는 그대로 유지되며, 여성과 아이를 타겟으로 한 블로그나 트위터, 페이스 북 등을 추가적으로 활용

다. 성과 및 시사점

❷ 키워드 검색의 변화

- ‘멍 빨리 없애는 법’의 키워드는 전년 동기 대비 33% 감소한 반면, 베노플러스겔은 557% 증가 → ‘멍 빨리 없애는 법’으로 검색했던 사람들이 베노플러스겔로 바꾸어 검색했다는 것을 의미

그림 50 ● 키워드 변화에 따른 매출 성장 그래프



자료 : (주)유유제약, 2012

- 베노플러스겔과 같이 제약 마케팅에 기여할 뿐만 아니라, 민간요법과 관련한 잘못된 소비자의 인식을 개선할 수 있는 가능성을 제시

02

SSD 온라인 버즈 마케팅

- 삼성전자

가. 추진 배경

- ❶ 기존 Small Data 분석에 근거한 정확성 결여의 문제는 비즈니스 적용에의 한계를 드러냈으나, 빅데이터 활용이 가능해 지면서 다양한 데이터나 니즈에 대한 정확한 분석과 더불어 소비자 인식이나 소비성향에 대한 예측이 가능해짐
- ❷ SSD 소비자 인식 마케팅의 목적

- Solid State Drive (SSD)는 하드디스크(HDD)와 비슷하지만, 기계적 장치인 HDD와는 다르게 반도체를 이용하여 정보를 저장하고, 임의접근으로 탐색시간 없이 고속으로 데이터를 입출력할 수 있으면서도 기계적 지연이나 실패율이 현저히 적음
- SSD 시장규모 확대 및 삼성 SSD 브랜드 가치 제고
- 소비자에게 SSD를 무엇이라고 이야기할 것인지 방향을 결정하기 위해서 SSD에 대한 다양한 인식의 틀을 살펴보고, 이에 근거한 마케팅 수행이 필요

나. 추진 내용

- ❸ SSD 마케팅의 성공을 위해 5가지 하부요소의 분석을 통해 SSD에 대한 소비자 인식을 분석
- SSD에 대한 기존인식

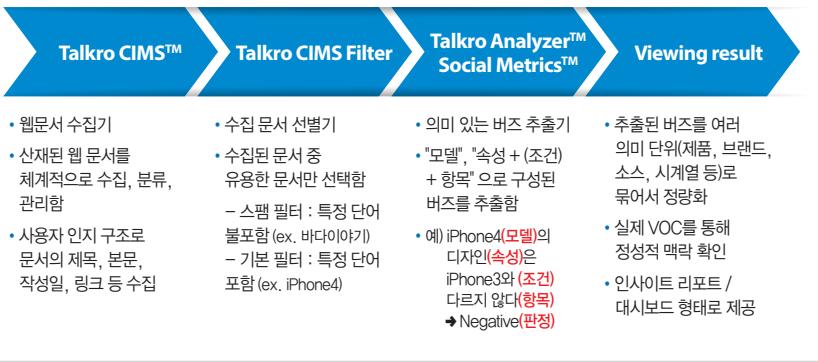
- SSD의 현재 상황
- SSD의 브랜드 조사
- 온라인 버즈(Buzz)의 테마
- 온라인 버즈의 출처

버즈 분석과정

- 인터넷의 자연어로 표현되어 있는 사용자의 의견을 수집하여, 고품질 자연어 처리 기술과 정교한 텍스트 마이닝 분석기술을 통해 제품 또는 서비스의 모델, 항목, 세부속성에 대한 평가 결과를 도출

그림 51 • Buzz Analysis의 전반적인 운영과정

❖ Process of buzz analysis



자료 : 삼성전자, 2013

- SSD 모델에 대한 실질적인 담론이 이루어지고 있는 전자제품 포럼과 하드웨어 포럼, 구매후기 위주의 쇼핑몰, 산업 뉴스를 다루는 미디어, 브랜드 사이트, 정부 기관 사이트 등을 대상으로 하여 모든 관련 웹사이트를 그 대상으로 함

그림 52 • Buzz Analysis 분석 대상

Site type

- Shopping Mall : 구매후기를 중심으로 제품구매 관련 질의응답도 이루어지는 사이트
- Forum : 특정 주제에 대한 이슈를 등록하면 Thread 형식으로 의견을 나누는 게시판의 형태가 주가 되는 사이트
- Electronics Portal : 주제가 전자기기에 한정되어 있는Portal
- Blog : 특정 주제에 대한 글을 포스트 형식으로 올리는 사이트
- Brand Site : 각 브랜드의 공식 사이트
- Media : 전자기기 및 IT 전반을 다루는 온라인 매체
- Organizations : Technology 관련 단체가 운영하는 사이트



자료 : 삼성전자, 2013

SSD 사용자와 사용처에 대한 요구사항 분석

그림 53 • SSD에 대한 인식

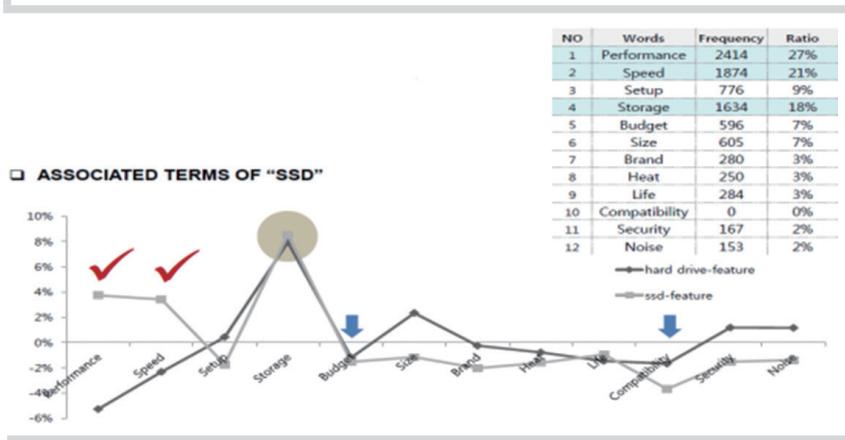
❖ IDENTIFICATION OF "SSD"

HDD 대체제 (replacement) HDD or SSD?	가격이 가장 큰 장애물
	소음과 발열 적음에 만족
	저장 기능이 우선시 되므로 큰 용량과 데이터 보존에 대한 안정성 기대
HDD 보조제 (co-exist) CPU, RAM or SSD?	PC hard user들의 인식
	저장 용도로 HDD, 부팅 용도로 SSD를 구분하여 사용
	컴퓨터 성능 향상을 기대하며 용량과 가격보다는 속도를 중요시
노트북용 HDD (비자발적 구매) High-end SSD or affordable SSD?	새로운 컴퓨터 구입시 일어남
	노트북 구매 시 탑재되어 나오는 경우가 빈번
	구매하는 노트북의 가격에 따라 SSD 선택 기준 나뉘어짐

자료 : 삼성전자, 2013

- SSD = Performance(CPU) + Speed(RAM) + Storage(HDD)
- SSD는 저장공간(storage)이 가장 중요하다는 점에서 HDD와 유사하지만, HDD와 달리 성능(Performance), 속도(Speed)가 중요하다는 점에서 컴퓨터 성능 향상에 도움을 주는 CPU, RAM과 유사함

그림 54 • SSD와 연관된 용어



자료 : 삼성전자, 2013

- 사용자 리뷰에서와 질의응답에서는 가격이, 뉴스(News)에서는 성능과 사용 상황이, 전문가 리뷰(Expert Review)에서는 속도에 관한 담론이 주요하게 전개되는 것을 확인. 또한 일반 소비자들이 전문가 보다 SSD속도에는 관대하고 가격에는 엄격함을 찾을 수 있었음

다. 성과 및 시사점

- 담론의 출처와 상품의 고객이 어디에 있는지를 정확하게 파악하여, 타겟 커뮤니케이션(Targeted Communication)을 가능하게 함

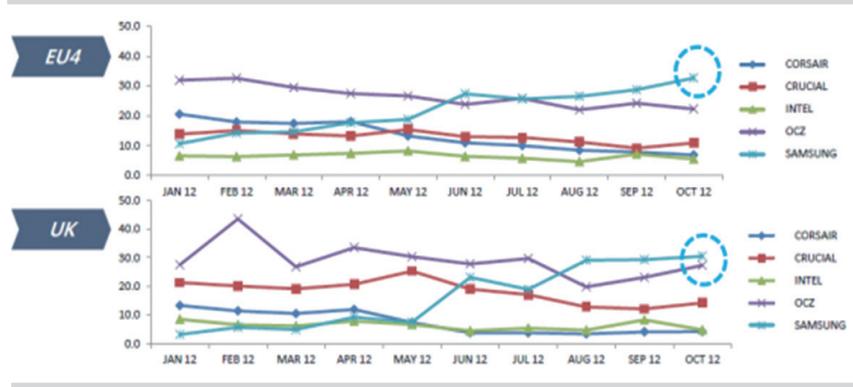
그림 55 • Targeted Communication 과정



자료 : 삼성전자, 2013

- 삼성전자는 SSD 시장의 후발 주자임에도 선두업체를 제치고 독일, 영국, 이태리, 한국에서 1위, 미국시장 2위의 성과를 기록

그림 56 ● 삼성전자의 SSD 시장 성과 추세 그래프



자료 : 삼성전자, 2013

03

C.S.I 프로젝트

– 신세계몰

가. 추진 배경

- 최근 3년간 연평균 40%의 성장을 달성한 신세계몰이 사이트 개편을 추진
 - 고객 데이터와 거래 데이터가 늘어나면서 영업부서가 원하는 데이터를 분석하려면 수 일이 소요
 - 온라인 쇼핑몰은 매우 빠르게 상품이 배치되고 새로운 캠페인이 생겨나는데 분석 결과를 얻기 전에 영업 담당자들이 ‘감(感)’에 의존해 캠페인 계획을 수립하고 마케팅 프로모션을 진행
 - 캠페인 성공률 분석 결과 지연에 따른 차후의 캠페인의 소비자의 요구를 반영하기에 애로 사항이 존재
 - M&A와 대형 백화점의 온라인 쇼핑몰 운영과 더불어 온라인 쇼핑시장의 경쟁이 과열
 - 빠른 성장세에도 불구하고, 다양해져 가는 소비자의 욕구를 충족시키기 위한 지속 적이고도 발빠른 대응이 요구됨
 - 개인 맞춤 서비스 트렌드에 부합하는 빅데이터 시스템 구축의 필요성

나. 추진 내용

- 서비스 개편에 있어 개인화 서비스를 위해 DW(Data Warehouse)와 CRM (Customer Relationship Management)를 구축하는 C.S.I (Customer Service Insight) 프로젝트를 추진

❷ 기존의 DW와 차별되는 온라인 고객을 위한 DW 구축

- 이전의 신세계몰은 신세계백화점이 운영하는 DW에 기반하여, 데이터마트를 구축해 사용하는 정도
- 신세계몰 고객과의 거래량의 증가에 따른 데이터마트의 한계점을 보완하고, 고객 개인화 맞춤 서비스 촉진을 위한 DW와 CRM 검토
- 온라인 몰 기능강화의 일환

그림 57 • 고객서비스 강화를 위한 신세계몰의 사이트 개편



자료 : IBM, 2012

❸ 실시간 분석 시스템

- 사용자가 원하는 관점으로 실시간 분석이 가능한 시스템을 구축하기 위해 IBM의 네티자 DW를 도입
- 신세계몰의 향후 성장성을 고려해 대용량 데이터를 빠르게 처리할 수 있는 최적의 하드웨어 소프트웨어 일체형 솔루션인 네티자를 벤치마킹테스트를 통해 선정
- 대용량 데이터 분석을 기반으로 하며, 정형 및 비정형 데이터의 분석이 가능해졌을 뿐만 아니라 신속성까지 갖춤

- 개인 고객 맞춤 캠페인을 위한 기반마련

그림 58 • 해피 바이러스 캠페인 분석을 통한 고객 선호 채널 파악



자료 : IBM, 2012

다. 성과 및 시사점

- 업무의 효율성과 성과달성을 실시간으로 모니터 할 수 있게 되면서, IT부서의 데이터 분석에 대한 시간과 노력이 절감
 - 고객 니즈의 실시간 분석이 가능해 짐에 따라, 빅데이터 활용을 통한 고객의 재구매율을 높이고, 새로운 고객창출이 가능
 - 온라인 쇼핑의 수요가 급격하게 증대되는 반면, 온라인 시장의 경쟁이 과열화 되는 상황에서 본 시스템은 고객에 대한 발빠른 대응력을 가능하게 할 뿐만 아니라, 궁극적으로 신세계몰의 경쟁력 제고에 기여

04

고객 맞춤형 서비스를 위한 빅데이터 센터

- 신한카드

가. 추진 배경

❶ 국민의 신용카드 사용이 상용화됨에 따라 카드업계의 경쟁이 과열되는 상황

- 신한카드와 국민카드가 올해 3,4분기 전체 카드 이용액이 각각 96조8000억원, 68조원으로 업계 1, 2위이지만, 같은 기간 체크카드 이용액은 KB국민카드가 14조1000억원, 신한이 11조3000억원을 기록하면서 국민카드가 신한을 다소 앞선 상태 (파이낸셜뉴스, 2013.12. 27)
- 카드사들은 새로운 서비스를 제공하는 신카드 출시를 통해 시장점유율 향상에 종력
- 각 카드사는 마케팅전략의 핵심인 매출액 증대를 위해 다각적인 노력을 기울이고 있는데, 신한카드가 빅데이터 이용을 통한 소비패턴 분석에 주력을 둔다면, 국민카드는 다양한 회원층을 확보하는 데 적극적임

나. 추진내용

❷ 빅데이터 센터 구축

- 2200만 고객의 정보를 모아 새로운 가치를 재창출해 고객에게 되돌려 주는 게 신한카드가 지향하는 빅데이터
- 기존에 다수를 대상으로 하던 매스마케팅을 벗어나 개별적인 특성을 강조한 마케팅에 관심을 기울이고 빅데이터센터를 새로운 성장동력으로서 추진

- 빅데이터센터를 영업 부문이 아니라 경영기획 부문에 배치해 단순한 영업이 아니라 상품개발, 영업, 상담 등 전사적 관점에서 활용하는 ‘빅데이터 경영’을 추진
- 빅데이터 센터는 빅데이터 기반 ‘마케팅 파트’, 내부 효율화를 위한 ‘비즈니스 인사이트 파트’, ‘빅데이터 플랫폼 파트’ 등 3개 부문으로 구성

❷ 빅데이터를 활용한 소비 패턴 분석

- ‘빅투그레이트(Big to Great)’라는 광고 문구에 ‘시장점유율 1위 카드사’라는 멘트를 전면에 내세운 광고를 시작한 가운데 최근 빅데이터를 활용한 ‘콤보(Combo) 서비스’를 확대
- 콤보서비스는 빅데이터를 통해 카드 회원들의 소비 성향을 분석한 결과를 바탕으로 선호도가 가장 높은 품목에 대해 할인 혜택을 집중 제공하는 서비스

그림 59 • 콤보서비스

제품명	서비스 Fee
카페	9,900원
베이커리	9,900원
영화	19,900원
골프	24,900원
도너츠	7,900원
레스토랑	23,000원
간 와인	8,500원
피자	22,000원
아이스크림	9,900원

자료 : 신한카드 홈페이지

- 2013년 12월부터 콤보 서비스 적용 카드가 하이포인트 카드, 빅플러스 카드 등 7종의 신용카드에서 러브 카드와 에스초이스 카드 등 2종류의 체크카드에도 추가로 적용돼 활용 폭이 확대 (대상 카드 수가 30만장에서 1700만장으로 획기적으로 증가)

다. 성과 및 시사점

- 개개인의 카드소비 성향패턴을 분석함으로써 고객의 만족도를 높임과 동시에 카드사의 매출액 증대에 기여
- 과거의 단순 매스마케팅(다수에 대한 마케팅)을 넘어 개개인의 니즈를 위한 고객과 기업의 소통을 활성화 시킬 수 있음
- 고객 개인의 소비패턴 분석은 단순한 소비 성향분석을 넘어 관련 분야 마케팅 전략에 활용 가능 (예 : 소비자에 인기있는 음식점에 대한 정보 제공)

I. 빅데이터 활용 시례

II. 빅데이터 활용 시례 Part 1.
미래장조회회부 사업사업 00종

III. 빅데이터 활용 시례 Part 2.
공공서적 11종

IV. 빅데이터 활용 시례 Part 3.
인간서적 13종

05

소비자 맞춤형 수요 매칭

– FNC 코오롱

가. 추진 배경

❶ 패션 산업의 특성상 새로운 방식 필요

- 제조업과 유통업이 결합된 패션 산업의 특징 및 감성 · 트렌드 · 디자인 · 영감 · 직관 등 사업 추진에 있어 주관적 영역이 많이 개입되어 정확한 예측을 하기 어려운 부분을 빅데이터 분석으로 해결 가능

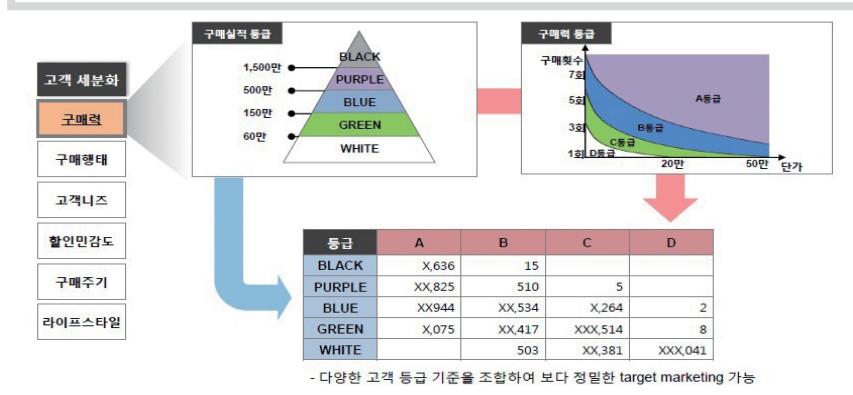
나. 추진 내용

❷ 고객 구매 데이터 및 코오롱 온라인 쇼핑몰 방문 로그 데이터 등 자체 데이터를 활용하여 분석

❸ 고객 가치 분석

- 매출을 기준으로 한 구매실적에 기반한 우수고객 관리 → 구매횟수와 단가 기준의 구매력 관리
- 구매력 등급 세분화 작업 완료

그림 60 • FNC 코오롱의 고객 가치 분석



자료 : Business Impact & Big Data 2014

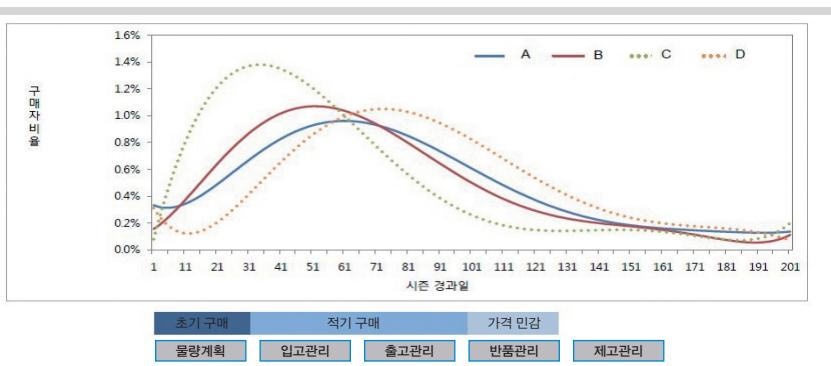
④ 고객 자산 분석

- 특정 브랜드의 매출을 고객의 분류와 구매단계를 통해 시작화
- 고객자산 성장모형으로 브랜드의 상태 예측

⑤ 브랜드별 판매 시점 분석

- 고객의 구매성향 분석 연동 프로모션 대상 설정 / 브랜드별 판매량 예측 기준 세분화
- 브랜드별 판매 주제 응용, 물류 물량 관리 효율화

그림 61 • FNC 코오롱의 브랜드별 판매 시점 분석



자료 : Business Impact & Big Data 2014

❸ 매장별 입점 고객 분석

- 성별, 연령별 매장 입점고객 데이터 확보를 위한 객수정보 시스템 도입
- 매장 단위의 매출 증감에 대한 원인 파악, 판매 효율 분석 기준 자료

❹ 온/오프 매출 상관관계 분석

- 온라인 고객 관심 상품과 오프라인 매장에서 판매 상관관계 분석
- 단기 판매 예측, 판매증가 예상 아이템 추출

❺ 패턴 분석을 통한 예측

- 과거 판매 패턴 분석을 통한 아이템별 시즌 총 판매량 예측
- 단기 판매예측 데이터와 연동, 대형 프로모션에 대한 예측치 감가 기준 함수 도출

다. 추진 성과

❶ 2013년 4분기 영업이익 12.7% 증가

- 극심한 경기불황의 영향으로 2013년 전체 매출액과 이익은 감소했지만, 패션 분야의 분기 최고 이익 달성을 위해 4분기 연결기준 매출액 1조 4253억, 영업 이익 689억으로 각각 전년 동기대비 0.7%, 12.7% 증가한 것으로 잠정 집계 (순이익은 311억원으로 77.7% 증가)

06

고객 맞춤형 매장 배치

- 현대백화점

가. 추진 배경

- 백화점이 분석할 만한 빅데이터를 갖지 못한 이유는 현재 관련 업계에서 운영되는 IT 시스템에서 기인하는데, 이들 IT 시스템이 고객관련 데이터를 전혀 수집하지 않기 때문
- 국내 백화점의 경우, 점포판매시스템(POS)으로 결제되는 데이터 가운데 실제 구매액수만 수집할 뿐, 고객의 특성이나, 어떤 매장이, 어떤 상품을 얼마나 판매하는 가에 대해서는 크게 관심을 두지 않음
- 백화점 관련 상품 유통에 있어서 고객과 구매상품에 대한 정보가 동시에 활용 · 분석 가능해진다면, 더 세밀한 고객 마케팅이 가능

나. 추진내용

● 빅데이터 구축 전략

- 고객관계관리(CRM)를 위한 수집 정보의 폭을 늘리고, 분류 방식도 단순 매출액 기준이 아니라 지역별 매출성향 혹은 실수익 비중, 빈도수
- 백화점 내 모든 시스템에서 데이터들이 연계돼 활용될 수 있도록 업무와 정보의 연동성, 통합성을 강화하며, 업무 부서간 협업 체제를 진행

- 고객이 구입한 제품의 종류, 사이즈 · 색상 등 POS 정보를 백화점 차원의 데이터로 축적하기 위한 연계 작업도 수행

❷ 고객 맞춤형 매장 배치

- 다각적인 빅데이터 활용을 위한 노력의 일환으로 고객의 소비패턴 분석에 따른 종체적인 매장 배치를 새로이 함으로 인해 매출이익 증대의 기대효과
- 빅데이터 분석을 통해 특정 상품을 원하는 소비자에게 연관된 상품을 제시하거나, 2차, 3차의 추가적인 구매를 유도

그림 62 • 고객 맞춤형 매장 배치



자료 : HSAd Webzine, 2012

❸ 빅데이터 분석을 통한 사례

- 현대백화점 본사 영업전략실 산하 영업기획팀은 고객데이터 분석을 통해 눈이 오면 식품 판매가 늘어나고 덩달아 아동용품 소비도 늘어난다는 추세를 발견
- 눈이 많이 오면 본인 물건보다는 생식품이나 아이들 물건을 먼저 사는 주부가 많아 해당 상품을 한 자리에 배치하는 마케팅 전략을 도입

다. 성과 및 시사점 도출

- 무차별 대량 DM 발송에 의한 비용 낭비와 비효율성을 지양하고, 고객별 맞춤 정보를 정확히 제공함으로써 매출을 예측
- 고객과 매출 정보에 데이터 마이닝을 실시, 패턴을 분석함으로써 성공률 높은 유형별 타깃 마케팅 가능
- 빅데이터 분석기술이 성숙하면 실제 니즈를 정확히 총족시키고, 추가구매를 더 많이 유도할 수 있는 다양한 추천 알고리즘이 가능해짐

07

맞춤형 고객 상품
추천 서비스

– GS샵

가. 추진 배경

- 온라인 커머스 시장에서의 경쟁력 강화를 위한 상품 추천 서비스 플랫폼 구축 요구
- 내부 역량을 활용해 다양한 서비스에도 확산시킬 수 있도록 자체 역량 확보가 우선시
- 아마존닷컴이 빅데이터 기반 추천 시스템을 통해 매출의 35%가 증가

나. 추진 내용

- 아마존닷컴의 추천시스템을 벤치마킹
 - 아마존의 추천시스템은 빅데이터 분석을 기반으로 하여 아마존 사이트에 접속한 사용자들이 남긴 로그데이터나 상품과 상품의 유사도 분석 등의 대량의 데이터를 분석해 상품을 추천
 - 누군가 책을 사려고 아마존에 접속하면 그의 이전 구매 내역 및 행동 패턴을 분석해 관심을 가질만한 책을 추천
 - 고객이 셔츠를 사면 그 셔츠와 어울릴만한 바지를 보여주며 함께 구매하도록 유도
 - 추천시스템의 성과를 확인한 아마존은 매년 이익의 10%를 추천 시스템 성능 향상에 투자

❸ 맞춤형 고객 상품 추천 서비스

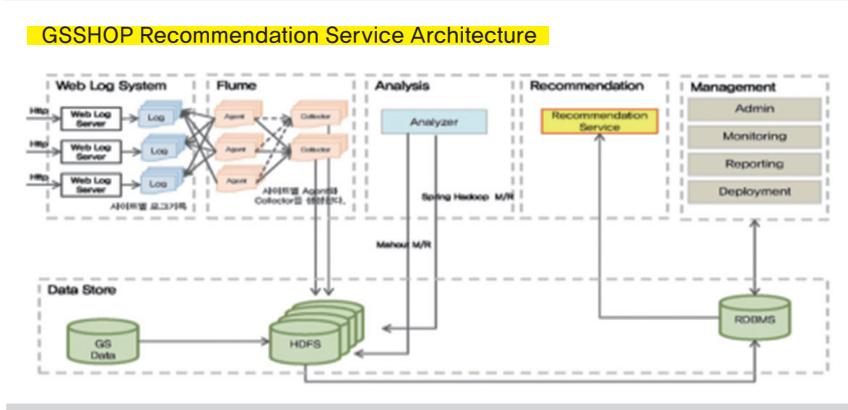
그림 63 • 노트북과 태블릿, 스마트 PC로 보는 GS샵



자료 : 공개SW포털, 2012

- 2012년 7월부터 하둡 기반 빅데이터 플랫폼으로 사용자의 구매 내역과 온라인 게시판에 남긴 텍스트 데이터를 이관하는 동시에 CEP(Continuous Event Processing)나 오피니언 마이닝(Opinion Mining) 등과 같은 빅데이터 관련 내부 연구 과제를 마련해 국내외 파트너사들과 함께 수행
- 국내 유통업체 최초로 하둡기반 빅데이터 플랫폼을 구축하고, 이를 기반으로 추천 시스템(사용자가 온라인 쇼핑 사이트에 접속하면 그가 관심가질 만한 상품을 추천)을 구축 (하둡에코 시스템 도입)
- 사용자들의 구매 내역을 바탕으로 상품 선택 패턴을 분석한 다음 그 결과를 토대로 사용자가 관심을 가질 만한 신제품이나 관련 상품들을 맞춤식으로 추천해주는 새로운 상품 추천 서비스를 구축

그림 64 • GS샵 상품 추천 서비스 아키텍처



자료 : 공개SW포털, 2012

- GS샵은 하둡 기반 시스템이 현재는 상품 추천 서비스에만 도입됐지만 보다 다양한 서비스로까지 그 적용범위를 확대해 나갈 방침

다. 성과 및 시사점

- 다양한 물에서도 사용 가능한 빅데이터 플랫폼 마련
- 내부 인력의 연구개발과 실제 적용을 병행할 수 있는 수준의 내부 개발자들 역량 확보
- GS샵이 구축한 추천 서비스를 위한 데이터 플랫폼은 좀더 거시적인 관점에서 해석하면 상품의 기획 및 준비 단계에 도움을 줄 수 있고, 단기적으로는 보다 정교한 실시간 상품 추천이 가능

08

맞춤형 여행 컨설팅 서비스

- 롯데관광

가. 추진 배경

- 빠르게 변하는 여행산업의 고객요구와 시장변화에 유연하게 대처하기 위해 업무 프로세스의 효율적 개선 및 통합이 필요하다는 문제점을 인식
- 기존 업무 시스템은 각 부문에 대한 연계가 부족해 업무의 효율성이 떨어졌으며, 정산 및 반제 시스템의 연계부족으로 일괄적 업무 처리에 애로사항이 존재
- 업무와 관련한 정보의 표준 부재로 통계의 어려움, 고객정보 오류 및 누락으로 인한 효율적 고객관리의 어려움

나. 추진 내용

❶ 빅데이터를 통한 고객 관리

- 소비자 행동 분석을 통한 고객관계관리 부문을 빅데이터 형태로 구축해 운영업무를 효율화하고 고객별로 차별화한 영업전략
- 표준화된 시스템 구현으로 업무효율을 극대화하고, 효율적 고객관리를 통한 비용 절감을 위한 고객관리 및 경영지원시스템을 구축
- 고객통합데이터베이스를 운영해 고객을 분석하고, 마케팅 전략 수립 · 고객접점 채널 통합 · 목표 고객 표적 마케팅을 실시하는 소비자 행동분석 · 마케팅 시스템을 구축

- 수집한 고객정보를 표준 정보로 축적해 고객에 대한 분석을 심도 있게 할 수 있게 됐으며, 신규고객 창출 뿐 아니라 기존 고객의 재구매율을 높이는데 활용
- 고객의 특성에 맞도록 상품을 분류하고, 각 고객이 민감해하는 부분을 가격·할인·항공사·숙박·주말이용 등으로 구분해 최적의 상품을 제시할 수 있는 환경을 구축

그림 65 • 롯데관광 홈페이지

The screenshot displays the homepage of Lotte Tourism's website. It features a top banner for 'Turkey is Good' with hot air balloons and Korean text. Below it are sections for international flights (e.g., GS Hubs, Japan, China), domestic flights (e.g., Seoul, Daegu), and travel packages (e.g., Seoul-Pusan, Seoul-Gyeonggi). A 'Travel Fair' section highlights '100% Travel Points'. On the right, there's a 'Hot Hotel Booking' section with dropdown menus for destination, date, and room type, and a 'HOT sale' banner for a lion statue in Busan.

자료 : 롯데관광 홈페이지

※ 빅데이터 분석을 통한 예

- 해외경험이 많은 고객일수록 일본, 유럽을 선호하는 경향이 높음. 결혼기념일 선물로 부부동반 여행을 많이 가는 경향이 있음

다. 성과 및 시사점

- 고객 중심 업무환경 구축과 시장의 변화에 반응하는 시스템 구축
- 어떤 상품을 제시했을 때 고객들의 반응이 높은지 고객층별로 구분이 가능

09

자동차 고객관계 관리 프로젝트

- 현대자동차

가. 추진 배경

- 최근 사회 전반에 빅데이터를 활용한 마케팅이 활성화되는 추세에 있고 빅데이터 분석을 통한 마케팅의 긍정적 효과가 검증되면서, 자동차 시장에서도 빅데이터 분석 사업이 본격화되고 있음
- 빅데이터의 활용을 통해 디지털 환경에서 만들어지는 방대한 자료를 분석하려는 사회 현상이 나타나면서 새로운 시각이나 법칙을 발견할 가능성이 증대
- 세계적인 자동차 기업들이 현재 빅데이터 분석사업에 관심을 가지면서 시장 경쟁이 한층 강화될 것으로 예측
- 현대자동차는 차량의 품질 · 서비스 향상을 위한 노력의 일환으로 빅데이터 분석 사업의 필요성이 증대

나. 추진 내용

● 사업자 선정과 플랜

- 사업자로서 테라데이터를 선정하여, 2014년 4월에 프로젝트를 마무리 할 예정
- 데이터 분석 프로세스 정립과 분석 모델 개발을 바탕으로 향후 인프라 구축 등 본 사업에 착수할 예정

❷ VCRM¹⁵⁾ 데이터 프로젝트의 목표와 내용

- VCRM 데이터 분석 프로젝트는 차량에서 생산되는 각종 대량 데이터를 체계적으로 관리
- 실시간으로 쌓이는 다양한 정보를 빅데이터 기술을 활용해 분석·활용하기 위한 체계를 갖추는 사업
- 데이터 분석을 바탕으로 제조 프로세스 개선을 통한 생산성 제고, 소비자 서비스 역량 강화, 새로운 비즈니스 창출을 제고
- 제조나 폐기 단계에서 축적한 데이터로 생산효율·재활용률 제고도 가능

그림 66 • VCRM 데이터 분석 프로젝트



자료 : Business IT Consulting Forum, 2013.12.19.

❸ VCRM 도입에 따른 활용 가능성의 예

- 특정지역에서 준중형차 검색 비중이 늘어날 경우, 해당지역에서의 준중형차 마케팅을 강화

15) 자동차고객관계관리(VCRM : Vehicle Customer Relationship Management)

- 특정 제품을 구입한 사람들이 가장 많이 검색한 검색어를 분석해 소비자들에게 적용할 마케팅 기법을 개발
- SUV를 구입하는 고객의 경우 레저를 고려하는 경우가 많다는 일례를 기반으로 SUV와 레저용품을 묶어 마케팅 전략을 수립
- 특정모델의 자동차와 아기용품을 묶어 마케팅하는 점도 빅데이터 활용의 사례로 이용될 수 있는데, 실제 수입차 주력제품을 구입하는 구매자가 20~30대의 신혼 부부나 자녀가 어리다는 점에 착안하여 아기용품을 함께 마케팅한 결과 그 효과가 상당했다는 점에 착안하여 다양한 마케팅 전략의 수립이 가능
- 센서 데이터 분석으로 판매 차량이 문제없이 운행 되고 있는지, 부품에 이상이 없는지 등을 사전에 파악해 대량 리콜 사태 등을 예방
- 특정 지역에서 운행되는 차량이 공통적으로 어떤 부품을 자주 교체하는지 등을 파악해 서비스 역량을 제고

다. 성과 및 시사점 도출

- ❶ 빅데이터 분석을 활용하여 다양한 소비자의 니즈를 반영하는 특화된 마케팅
- ❷ 고객이 제조사에 대한 문제점을 제기하기 전에 제조사가 문제점을 파악해 고객에 다가갈 수 있고, 실제 어떠한 고민을 운전자들이 가지고 있는지 파악하여 선제적 대응을 가능하게 함

10

스마트 미디어 솔루션 서비스

– 삼성전자

가. 추진 배경

- 스마트폰 사업과의 시너지를, 나아가 스마트TV 등 각종 스마트 기기들과의 동반 성장을 추진하는 것이 목표
- 방대한 데이터를 한데 수집하고 그 속에서 의미 있는 정보를 찾아내야 하기 때문에 빅데이터 기술이 필요
- 2013년 초반까지도 삼성전자는 빅데이터 분석시스템과 관련하여 외부에 의뢰, 일부 마케팅 업무에 활용하는 정도에 그쳤으나, 최근 자체적인 빅데이터 시스템 구축의 필요성을 절감
- 최근 구글, 애플 등 IT업계 공룡들이 대규모 데이터센터를 설립하며 차기전쟁을 준비하기 시작

나. 추진내용

- 본격적인 빅데이터 분석 구축
 - 삼성전자는 최근 미디어솔루션센터(MSC) 내 빅데이터 전담 조직을 신설
 - 지난 6월 삼성전자는 자사 빅데이터 표준 시스템에 오라클을 선정
 - 빅데이터가 마케팅이나 고객 관리뿐만 아니라 반도체 설계, 데이터 분석 등 다양한 영역에서의 활용 가능성

그림 67 • 빅데이터 분석 시설 구축



자료 : SAMSUNG TOMORROW

● 스마트 미디어 솔루션 서비스

- 사용자 로그(시간에 따라 정보가 기록)를 분석, 콘텐츠 이용 패턴과 선호도를 파악하여, 이를 신규 서비스 개발과 사업 모델에 반영하는 작업을 추진
- 현재 전 세계의 수억 명이 넘는 사람들이 사용하는 자사의 스마트폰이 데이터 축적에 새로운 기회를 제공
- 자사 스마트폰 사용자들의 동의 하에 다양한 정보를 직접 수집한 뒤 데이터를 분석해 맞춤 추천 정보 등을 제공

그림 68 • 스마트 미디어 솔루션 서비스

The screenshot shows the Samsung B2B homepage with a banner for 'Mobile Business Solution' and sections for 'Education solution' and 'Hotel Solution'. On the right, there's a sidebar with links to 'Samsung B2B 회원가입' and a contact number '무료문자 보내기 1588-3536'.

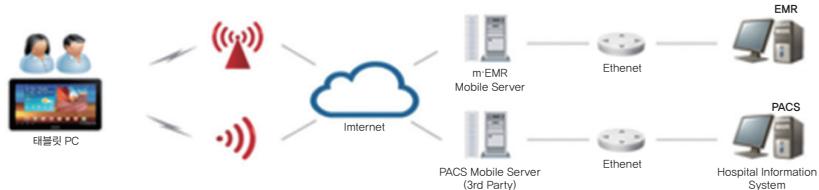
자료 : 삼성전자 B2B

❶ 솔루션 서비스의 예

- ① 모바일 EMR (Electronic Medical Record) 솔루션
 - 실시간 의료 정보 제공으로 의료진과 환자의 만족도 제고
 - 삼성전자의 태블릿 PC를 통해 이동 중에도 의료정보를 확인할 수 있는 솔루션
 - 환자치료 및 의료진 업무효율 향상, 시각데이터로 직관적인 환자의 상태 파악, 환자의 만족도 증대

그림 69 • 모바일 EMR 솔루션

❖ 모바일 EMR 구조



❖ 모바일 시스템 예시

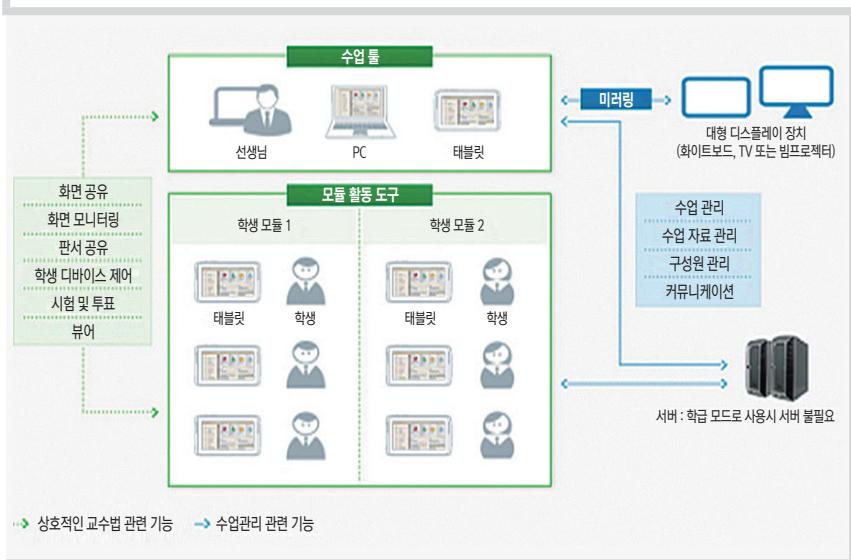


자료 : 삼성전자 B2B

② 삼성 스마트 스쿨

- 차세대 교실을 위한 통합 디지털 교육 솔루션으로 교사 태블릿, 교사 PC, 학생용 태블릿, 서버, 대형 디스플레이로 구성되어 교사와 학생 간 상호적인 교수법과 수업관리 기능을 제공
- 교사와 학생 간의 자유로운 소통을 가능하게 하고, 효과적인 멀티 미디어 콘텐츠 이용을 통한 역동적인 수업의 가능성

그림 70 • 삼성 스마트 스쿨 구성도



자료 : 삼성전자 B2B

다. 성과 및 시사점

- 빅데이터 센터를 통해 시장 분석 및 예측 역량을 강화
- 스마트 기기의 폭 넓은 활용을 가속화하고, 시너지 효과를 기대할 수 있으며, 관련 기기와의 동반 성장을 가능하게 함

11

모바일 빅데이터 맞춤형 서비스

– NHN(네이버)

가. 추진 배경

- 데이터는 의사결정뿐 아니라 검색 서비스의 바탕이 될 수 있다는 인식 하에, NHN은 검색엔진인 네이버 자체가 빅데이터임을 강조
- 네이버 이용자는 하루 2천만 개의 서로 다른 질문을 던지고 그에 대한 결과로 NHN은 130억개 문서에 순위를 매김
- 포털 사이트의 검색 기능이 점점 더 '이용자 맞춤형' 서비스로의 진화
- 기존의 검색 서비스가 기계적으로 질의어와 관련된 문서 정보만 제공하는 시대는 한계점을 드러내고 있음

나. 추진 내용

● 모바일에서 맞춤형 검색 서비스

- 이용자들의 검색창 활용방식에 주목해 지역 명소 추천, 응원메시지 등을 공유하는 기능 제공
- 이용자들이 모바일 검색에서 관련 이미지를 검색하려는 의도가 비교적 명확할 경우, 이미지 영역이 모바일 검색 결과의 가장 상단에 노출되는 방식

- PC와는 차별화되는 모바일 상의 검색어 군(群)에 대해서는 개별 키워드별로 검색 결과를 최적화시켜 제공한다는 계획
- 사용자들의 보다 섬세한 검색 니즈에 부합하기 위해 개별 검색어에 최적화된 검색 결과와 UI(User Interface)를 제공

❷ 모바일 검색 서비스의 예

① 지역명소

- 이용자는 지역명과 가볼만한 곳으로 검색을 하면 검색 결과 최상단에서 지역 명소를 찾을 수 있음
- ‘부산 가볼만한 곳’이라고 검색창에 입력하면 검색결과에 ‘광안리 해수욕장’, ‘베프광장’, ‘금강공원’등의 장소가 뜨게 됨

② 검색창을 활용해 응원메시지를 다른 친구들과 공유

- 모바일 검색창에 좋아하는 유명인에 대한 응원 메시지를 작성하면 해당 검색어의 입력 순위를 다른 이용자들과 함께 공유할 수 있는 서비스를 제공
- 팬의 입장에서 응원, 죽하 등의 긍정적 메시지가 오가게 되는 유명인, 웹툰작품, 프로야구구단 등의 키워드와 ‘축하해요’, ‘사랑해요’, ‘힘내요’등과 같이 긍정적 키워드가 조합된 검색어에 대해 적용
- 이 서비스는 이용자들이 질의어를 얼마나 입력했는지를 등수로 표현하는데, ‘윤후 사랑해’라고 입력하면 1만2천533등으로 검색했다고 검색 결과가 도출

그림 71 • 모바일 검색 서비스 예시



자료 : 네이버

다. 성과 및 시사점

- 빅데이터에 기반한 맞춤형 검색 서비스는 불필요한 검색을 줄여 원하는 정보를 더욱 손쉽게 찾을 수 있게 검색 기능을 고도화
- PC와는 차별화되는 모바일 상의 검색 결과 최적화를 통해 기기의 특성에 따른 고객의 요구사항 반영이 가능

12

스마트 통합물류 시스템

- CJ 대한통운

가. 추진 배경

- 물류기업에 있어 전문가들은 빅데이터를 '미래 물류시장의 신호을 읽어내는 안테나'로 요약
- 빅데이터 분석 시스템 구축을 위해 세계적인 물류기업들이 빨빠르게 대응

그림 72 • 빅데이터 인 로지스틱



자료 : DHL 홈페이지

- 최근 DHL이 물류산업에서의 빅데이터 활용방법을 소개한 '빅데이터 인 로지스틱스 (Big Data in Logistics)'를 발표

- 빅데이터가 물류 운송의 최적화 커다란 변화를 가져오고 있음
- 많은 물류 기업들은 이미 빅데이터의 활용능력이 추후 기업의 성패를 가늠할 핵심 경쟁력으로 인식

나. 추진 내용

● CJ 대한통운의 스마트 통합 단말기

- 스마트 통합 단말기에 디지털 운행기록계를 결합해 차량의 위치와 경로, 화물의 상태, 연료 소모량, 속도 등을 실시간으로 파악하고 관제하여 물류 차의 전반적인 관리와 통제를 효과적으로 할 수 있도록 함

그림 73 • CJ 대한통운의 스마트 통합 단말기 사용



자료 : CJ 대한통운

- 최적의 운송경로를 찾아 택배 배송기사들에게 지금된 스마트폰과 태블릿 등에 네비게이션 기능을 통해 전달
- 운전자가 화물을 내린 장소에서 가장 가까운 곳에 있는 다른 화물의 정보를 실시간으로 확인할 수 있어 공차율을 낮출 수 있을 것으로 전망

- 디지털 운행기록정보를 분석해 급출발, 급정지, 급가속 등의 10대 안전지표를 관리하여 각종 안전사고를 줄일 계획
- 전체 시스템을 새롭게 하여 정보 수집 및 저장 시스템도 단일화 함으로써 빅데이터를 활용하여 화물비용 관련 데이터를 한 곳에서 처리할 수 있도록 함
- 각종 차량의 수집된 정보는 빅데이터 분석을 통해 연료절감을 위한 노력을 꾀할 수 있고, 물류 효율성을 높일 수 있음
- 빅데이터 분석으로 불필요한 지출을 억제시켜 인력 및 비용절감 효과를 누릴 수 있음

통합물류 시스템 확장

- 최근 들어 증가한 택배 물량을 통합적이고 효율적으로 관리하기 위해 파워7+ 프로세서로 보다 강화된 성능을 제공하는 IBM 파워 시스템즈를 추가적으로 도입
- 통합물류시스템 확장을 통해 국내 택배 물동량을 사전에 예측하고 선제적으로 관리함으로써 운영 효율성과 안정성을 증대
- 추석, 설날 등 택배 물량이 급증하는 성수기에 대비한 처리 용량을 확보해 갑작스러운 시장 변화에도 새로운 추가 업무에 유연하게 대응

다. 성과 및 시사점

- 통합 물류 시스템 확장을 통해 안정적이고 최적화된 IT 인프라를 마련하고 고객 서비스 경쟁력을 한층 더 강화
- 서비스의 속도 향상과 연속성을 보장하는 안정된 기반을 확대해 물류 시스템의 전반적인 퍼포먼스 상승은 물론 운영 및 관리 비용의 절감

13

언더라이팅의 예측모델 활용 방안

– 푸르덴셜생명보험

가. 추진 배경

- 빅데이터를 이용한 예측모델과 이를 통한 변화는 산업전반에서 나타나고 있음
- 미국의 경우 마케팅, 언더라이팅, 클레임, 리스크 등의 분야에서 예측 모델이 활용되고 있음. 약 40%의 보험사들이 예측모델을 매출에 활용해서 보다 판매의 확률이 높은 고객을 발굴
- 생명보험사 예측모델은 최초 가입시점부터 언더라이팅¹⁶⁾, 보전, 지급시점까지 전 부분에 걸쳐 적용될 수 있으며 응용이 가능함
- 특히 과거 및 현재의 정보를 기반으로 미래의 위험을 예측해야 하는 언더라이팅 입장에서 예측모델¹⁷⁾의 발전이 매우 중요해짐
- 기존의 언더라이팅 과정에서 우량체(건강체)¹⁸⁾는 BMI, 혈압, 니코틴 수치에만 의존하여 단순하게 판별됨. 그러나 피보험자의 Lifestyle이 더 많은 영향을 끼침
→ 보다 정확한 우량체를 선별할 수 있는 예측모델 도입이 필요해짐

16) 언더라이팅(Underwriting) : 생명보험 계약시 계약자가 작성한 청약서상의 고지의무 내용이나 건강진단 결과 등을 토대로 보험계약의 인수 여부를 판단하는 최종 심사 과정

17) 예측모델 : 미래의 특정 행동이나 사건을 예측하기 위한 통계학적 모델로 예측분석의 한 과정이며 예측 분석은 확률과 경향을 예측하기 위한 데이터 마이닝(Data mining)의 한 분야. 예측 모델은 미래의 행동이나 사건을 예측하는 예측인자와 변수들로 이루어져 있음

18) 우량체 : 일반고객에 비해 사망의 확률이 낮다고 판단되는 고객을 지칭. 보험회사에 이들을 대상으로 보험료 할인 등의 혜택을 주기 위해 언더라이팅을 진행

나. 추진 내용

- ❷ 언더라이팅에서 우량체 선별을 기준의 All or nothing이 아닌 스코어링 방식인 Debit & Credit 방식을 적용

- 스코어링 방식 : 개별 변수가 가진 리스크의 크기에 숫자를 매번 더하는 방식으로 산정하는 방식. 일반적으로 높은 점수는 더 위험하고 낮은 점수는 덜 위험한 것으로 판단함
- Dbit/Credit 방식 : 항목별 각각의 점수를 부여 후 미리 정해진 가중치에 따라 점수를 계산하는 것

그림 74 • 생명보험사 예측모델의 적용



자료 : 푸르덴셜생명보험 계약심사팀, 2013

- 과거 근 10년 이상의 데이터를 바탕으로 고객집단별 평균 사망률을 추정 및 스코어 점수로 구현
- 이러한 내용을 수행하기 위해 일반화 선형모형(Logistic) 및 일반화 가법모형(GAM)에 대한 이론적 배경과 사례분석을 통해 모형을 추정하고 해석을 시도함

- 언더라이팅 실무에서는 이를 스코어 테이블로서 개발하여 특정 스코어 기준치 이하일 시 우량체를 적용함

그림 75 • 우량체 판별 스코어링 예시

	연령	SGOT	r-GTP	HBs4g	Hbe4g	FB5	영업보험료	니코틴	수축기혈압	일사기입금액	BMI	혈뇨	Hemoglobin	계
시나리오 1	38	12	31	Neg	Neg	65	28,300	Neg	120	3000만	20	Neg	10.8	
Score	-13	19	-27	-6	-18	-3	9	-24	-7	-2	-9	-3	44	-40
시나리오 2	54	45	87	Neg	Neg	124	125,800	Pos	150	5000만	28	Neg	12.5	
Score	19	56	-7	-6	-18	42	5	3	35	2	11	-3	4	143
시나리오 3	30	32	51	Neg	Neg	102	48,500	Pos	135	1억	28	Neg	11.7	
Score	-21	28	-23	-6	-18	2	9	3	3	5	11	-3	4	-6

자료 : 푸르덴셜생명보험 계약심사팀, 2013

다. 성과 및 시사점

● 모형적용 결과 우량체 선별 시 피보험자의 의적 건강상태 보다는 생활습관 등의 환경적 영향이 중요하다는 것을 알게됨

- 간수치, 연령, 혈당, 보험료 보다는 피보험자의 연령/직업/소득 등에 따른 음주 및 흡연습관, 신체활동 등 라이프 스타일이 우량체 선별에 중요
- 이를 바탕으로 숫자사정 우량체평가 테이블을 개발하거나 경험사망율과 비교한다면, 더 적절한 방법으로 우량체를 분류할 수 있을 것으로 예상됨

● 추가적으로 정확한 모델의 개발을 위해서는 여러 통계적 기법을 통해 지속적으로 모델을 유지 보수해 나가는 과정이 필요함

참고문헌

디지털데일리(2013.3.19.), 아마존 빅데이터 성공모델 벤치마킹한 GS홈쇼핑 … 성과는?

디지털 타임즈(2013.11.5.), 빅데이터로 고객관리 ‘롯데관광’

매경이코노미(2012.9.24.), [돈이 되는 정보 ‘빅데이터’]국내 기업, 빅데이터 활용은 … 아직은 걸음마 단계

매일경제뉴스(2013.12.4.), 신한카드, 카드업계 최초 ‘빅데이터 센터’ 출범

매일경제뉴스(2013.12.19.), 마네킹이 고객정보 수집 … 맞춤마케팅 대박 나이 · 성별 · 인종 정보 DB에 차곡

빅데이터(2013.1.28.), 네이버, 빅데이터 분석한 맞춤형 모바일 검색 확대

머니투데이(2014.2.13.), 코오롱인더, 작년 4Q 영업익 12.7% 증가

전자신문(2013.8.16.), 신세계, 빅데이터 도입해 온라인몰 역량 강화한다

(주)유유제약(2012.12.), 빅데이터 분석을 통한 베노플러스겔 마케팅 활용사례

최준용, 최항식, 유승현, 최한승(2013), 언더라이팅의 예측모델 활용방안 : 우량체 예측모델을 중심으로.
생명보험언더라이팅 학술논문

파이낸셜뉴스(2013.12.4.), 카드업계, 빅데이터 경쟁 본격화

파이낸셜뉴스(2013.12.27.), 신한 · KB국민카드 업계 1위 경쟁 치열

C3R 한국클라우드컴퓨티연구조합(2013.5.14.), 삼성전자, 빅데이터 도입 속도전 … SSP 적용 위한 ISP완료

한국패션협회(2013), CEO 뉴스레터(2013.02.23.~ 2013.02.26)

Auto Times(2013.8.16.), 자동차 마케팅에 빅데이터 도입한다면? Business Consulting

Chosun Biz(2013.12.2.), CJ대한통운 빅데이터 활용해 택배 화물차량 관리

Etnews(2013.6.12.), 삼성전자 ‘빅데이터, 스마트폰부터 먼저 접목’미디어솔루션센터 내 조직 신설

Etnews(2013.8.15.), 신세계, 빅데이터 도입해 온라인몰 역량 강화한다

Enews(2013.8.16.), 네이버 모바일에선 심심할 틈 없다?

Enews(2013.12.12.), 현대차, 빅데이터 분석 사업 본격화 … 글로벌 업체와 서비스 경쟁 스타트

Forum(2013.12.19.), 영림원소프트랩, SAP · 오라클 · MS와 경쟁 ‘대표 ERP기업’

HS Ad(2012.11.12.), Strategystudy 쇼핑의 과학 –매장 내에서의 심리와 선택행동

IBM(2012), 신세계몰, DW에서 빅데이터 분석까지 효과만점의 시스템 구축. IT World.

IT world(2013.07.22.), CJ대한통운, 물류 시스템 확장 위해 IBM 파워 7+도입

LG CNS(2013.1), Technology Inside : ‘빅데이터 플랫폼의 미래’, LG CNS R&D Journal

Open Source Software(2012), GS Shop – 하둡기반 빅데이터 분석 플랫폼

삼성전자(2013.01). ‘Online Buzz Analysis Report : Understanding SSD market and customers’

ZDNET Korea(2013.1.14.), 백화점이 VIP마케팅 밖에 못하는 이유

빅데이터 분석 활용센터 이용방법

온라인으로 빅데이터 분석 활용센터 홈페이지 (<http://www.kbig.kr>)에 접속하여 회원가입 후 자원 (HW/SW 등)을 할당 받아 사용



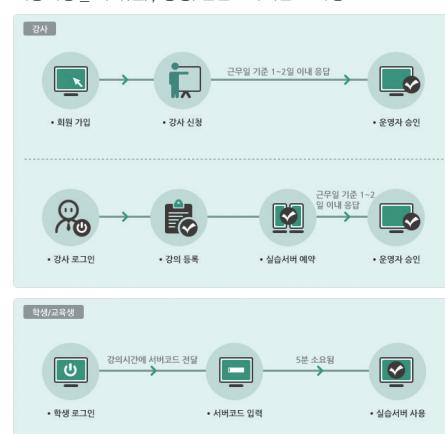
이용대상 _ 중소기업, 창업준비자, 커뮤니티 개발자 등

사업지원 인프라



이용대상 _ 대학(원), 공공/민간교육기관 교육생

교육지원 인프라



창조경제 실현을 위한
2013 빅데이터 국내 사례집

2014년 3월 인쇄

2014년 3월 발행

발행인 장 광 수

발행처 한국정보화진흥원 빅데이터 분석활용센터

집 필 신신애, 김성현, 송경빈, 류승희, 정규진, 송리라

주 소 서울시 종구 청계천로 14

연락처 (02) 2131-0114

인쇄 전우용사촌(주)

〈비매품〉

