

# 경상남도 빅데이터 활용 방안

연구책임  
박병주(경제산업연구실 연구위원)

2015. 12

경남발전연구원

# 요 약

## □ 경상남도 빅데이터 사업 추진 방향 설정

- 정부 3.0 등장 이후 합리적·과학적 행정을 구현하기 위해 필요한 우선적인 빅데이터 활용 분야 선정과 빅데이터 활용 방안 마련에 대한 경남도의 노력이 요구되어 본 연구는 경남도 주요 현안 중심으로 빅데이터 시범사업을 발굴하고, 경남도 빅데이터 활용 방안을 마련하는 것을 목적으로 함
- 경상남도 빅데이터 사업 시행의 목표는 빅데이터 활용체계 구축 및 정비, 빅데이터 전문성 확보, 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공으로 설정하고, 목표 달성을 통해 「Big Data 기반 도민 삶의 질 개선」이라는 비전 실현을 추진 방향으로 설정함

## □ 경상남도 빅데이터 시범사업안 제시 및 향후 과제

- 경상남도 빅데이터 시범사업은 복지와 안전 분야에 중점을 두어야 한다는 의견이 가장 많았고, 시범사업(안)은 복지(2), 안전(1), 관광(2), 기타(1)로 총 6개 사업을 제시함
- 시범사업 이후 경남의 빅데이터 체계를 보다 확고히 하기 위해서는 향후 지속적인 빅데이터 서비스 제공을 위한 기반(플랫폼, 전담조직, 하드웨어 및 소프트웨어, 기초데이터 등)이 구축되어야 하고, 이 기반을 통해 공공데이터 공유, 특정 분야 자료 실시간 공유, 빅데이터 분석 모델 확산, 빅데이터 분석 서비스 등이 제공되어야 할 것으로 판단됨

# 목 차

|                                    |    |
|------------------------------------|----|
| I. 연구배경 및 목적 .....                 | 1  |
| II. 국내·외 빅데이터 추진 사례 .....          | 3  |
| III. 경상남도 빅데이터 사업 및 체계 구축 방향 ..... | 14 |
| IV. 경상남도 빅데이터 시범사업 추진 방안 .....     | 31 |
| V. 경상남도 빅데이터 시범 분석 .....           | 66 |
| VI. 결론 .....                       | 72 |
| 참고문헌 .....                         | 73 |

## I

## 연구배경 및 목적

## 1. 연구 배경

## □ 빅데이터 분석 트렌드 급속 확산

- 빅데이터는 데이터를 구성하고 있는 물리적 하드웨어로부터 시작해 이를 기반으로 하는 애플리케이션과 소프트웨어를 포괄하는 거대 플랫폼으로 단순한 데이터의 크기가 아니라 데이터의 형식과 처리 속도 등을 함께 아우르는 개념으로 기존 방법으로는 데이터의 수집, 저장, 검색, 분석 등이 어려운 데이터를 총칭함
- 빅데이터는 우리 사회 전반적인 분야에서 핵심 키워드로 자리매김 하고 있고, IT 분야를 비롯해, 공공, 의료, 복지, 교통 등 사회 전 분야로 확산되고 있는 상황으로 우리생활에 매우 밀접하게 자리잡고 있는데 아마존의 '추천 상품' 등이 빅데이터 분석 결과를 활용하는 것임
- 글로벌 빅데이터 시장은 연평균 17%로 급성장하고 있고, 2020년 609억 달러로 성장할 것으로 전망되고 있으며, 또한 빅데이터 분석을 통한 정책 효율성 제고를 위해 국가적 차원에서 빅데이터 관련 정책 및 사업을 추진하고 있고, 각 지자체에서도 공공데이터를 기반으로 한 빅데이터 분석을 통해 각 지자체 맞춤형 정책을 발굴해 나가고 있음

## □ 빅데이터 활용 방안 마련 필요성

- 특히 정부 3.0 등장 이후 합리적·과학적 행정을 구현하기 위해 필요한 빅데이터 활용 방안 마련과 우선적인 빅데이터 활용 분야 선정에 대한 경남도의 노력이 필요한 시점임
- 경상남도는 산업, 관광, 복지 등 다양한 분야에 걸친 정책 발굴 및 사업을 추진하고 있으나, 이들 주요 현안에 부서별 대응으로는 문제 해결을 위한

정책을 제시하기 어려운 경우가 발생하고 있어 경남도는 그간 축적되어 있는 다양하고 방대한 공공데이터와 도민들의 소셜데이터를 기반으로 객관적이고 과학적인 분석결과를 도출하고, 이를 기반으로 지역별·분야별 맞춤형 정책 및 사업을 발굴·추진해야 할 필요성이 있음

## 2. 연구 목적

- 본 연구에서는 경남도의 주요 현안 문제에 경남도가 보유한 공공 데이터 등의 빅데이터를 활용하여 과학적이고 체계적인 행정을 구현하기 위해 경남도에 빅데이터를 적용할 수 있는 정책 분야와 빅데이터의 단계적 활용 방안을 제시하고자 함
  - 이를 위해 경남도 주요 현안 중심으로 빅데이터 분석 사업을 발굴하고, 빅데이터 분석을 토대로 정책을 개발할 수 있도록 경남도 빅데이터 분석 및 활용 방향을 마련함
  - 2016년 경남도는 빅데이터 시범사업 추진을 계획하고 있는데, 시범사업이 경남도정의 주요 현안과 경남도의 공공데이터, 관련 데이터 확보 및 분석 가능성, 사업 소요 예산 등을 고려하여 추진될 수 있도록 경남도 빅데이터 시범사업 발굴 및 추진 프로세스 수립과 추진 방안을 제시하는 것이 본 연구의 목적임
- 또한 내년 경상남도 빅데이터 시범사업 추진을 대비하여 시군에 축적된 공공데이터들 중 일부의 분석을 통해 정책적 활용방안을 찾아보기 위해 U-City 사업(양산시 경로당 복지시스템 구축)을 통해 축적된 공공데이터를 활용하여 빅데이터 시범분석을 실시함
  - 빅데이터 시범분석은 양산시 노인인구 분포, 노인인구 연령대별 건강 현황, 지역별 질환 현황 등을 파악하고, 양산시 소재 산업단지와 노인인구 건강과의 관계를 시각화 작업을 수행하고자 함
- 본 연구는 추후 경상남도 빅데이터 사업 추진 시 방향 설정, 단계별 추진 전략, 실효성 있는 결과 도출을 위한 기준 자료로서의 활용이 가능할 것임

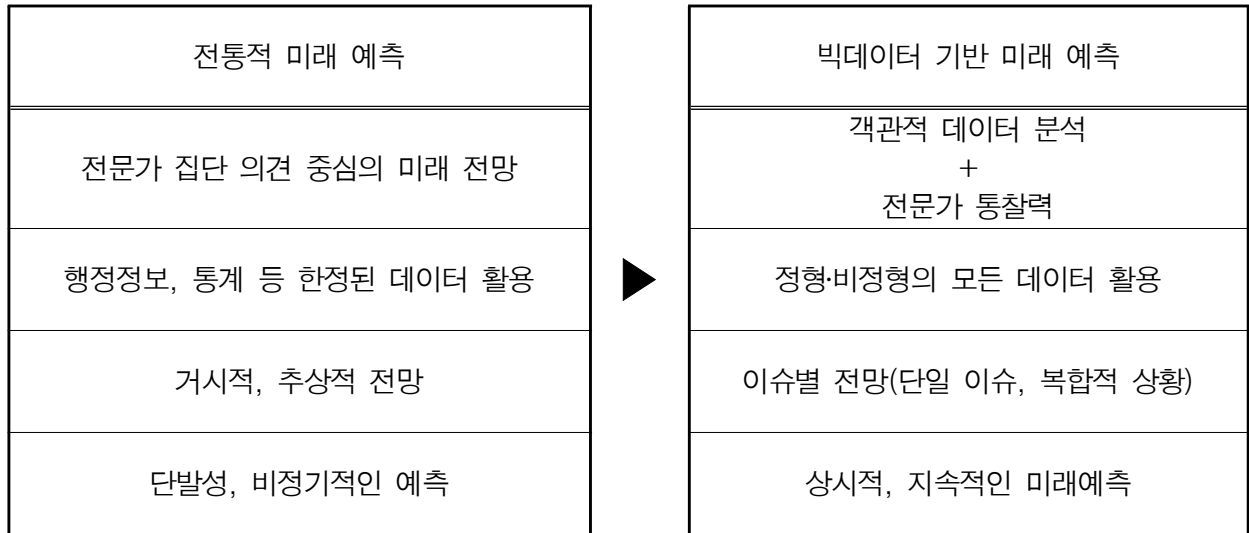
## II

## 국내·외 빅데이터 추진 사례

## 1. 빅데이터 분석의 필요성

- 빅데이터 활용을 위한 분석기법은 분석기술과 표현기술로 분류되는데, 빅데이터 분석 기술은 데이터 마이닝, 기계 학습, 자연 언어 처리, 패턴 인식 등의 기술에 기반하고 있고, 최근 비정형 데이터의 증가로 인해 자연 언어 처리 기술 기반 텍스트 마이닝 및 긍정과 부정을 의미하는 특정 언어 처리 기술 기반 오피니언 마이닝, 행위자 간 연결구조 분석에 기반한 소셜 네트워크 분석 그리고 유사한 특징을 가진 집단 도출을 위한 군집분석 등이 있음
  - 빅데이터 표현 기술은 분석 기술을 통해 분석된 데이터의 의미와 가치를 시각화하기 위한 기술로서 NodeXL과 Gephi, R 기반 패키지 등이 있음
- 빅데이터는 새로운 주요 자원으로 공공기관, 기업체 및 시민들의 필요에 따라 구성 및 맞춤형 시킬 수 있는 미래의 동적 자원임
  - 빅데이터는 행복하고 안전한 사회 구현을 위한 과거, 현재, 미래 분석 및 예측을 위한 자산이고 기업 및 공공분야의 더 나은 서비스를 제공하는데 있어 필요한 기준 제시 및 새로운 비즈니스 기회를 창출함
- 이는 문제해결 및 이슈 대응에 데이터 활용 중요도가 상승하고 있기 때문에 데이터 분석 기술의 발전은 급증하는 데이터에 효과적 대응이 가능하도록 하고 있음
- 이를 위해서는 빅데이터 자원 자체의 축적 및 관리뿐만 아니라 이를 가공하고 분석할 수 있는 차별화된 기술과 서비스가 필요함
- 이에 정부는 과학적 정책 수립을 국정운영 패러다임의 변화의 핵심으로 제시하고, 빅데이터를 활용한 체계적·과학적 문제 해결 대안 및 전략 마련을 위한 데이터 기반 미래전략 컨설팅 사업 기획하였고, 빅데이터 컨설팅 사업

은 '13년 빅데이터 사업화 지원을 위한 컨설팅을 추진하였으며 '14년도는 데이터기반 미래전략 수립을 위해 정확한 현안 진단에 기반한 의사결정 및 미래예측에 활용될 수 있는 다양한 모델 개발 및 어젠다 발굴에 중점을 두어 사업을 추진한 바 있음<sup>1)</sup>



자료)미래창조과학부, 빅데이터 기반 미래예측 및 전략수립 지원계획, 2014.

(그림1) 미래 전략 수립 및 예측 방법의 변화

## 2. 빅데이터 국내 동향

- 빅데이터와 관련해서는 행정자치부 공공정보정책과에서 빅데이터 정책 및 제도를 총괄하고 있고, 미래창조과학부 정보통신정책실 인터넷융합정책관 융합신산업과에서 2015년 8월 마련한 K-ICT 전략을 기반으로 사업을 추진하고 있음
- K-ICT 전략에서는 4개 분야 17개 과제를 중점 추진하는 것으로 계획하고 있는데, 이 중 SW·신산업 분야 9대 전략산업에 빅데이터가 포함되어 있음<sup>2)</sup>
  - 제조, 건강, 기상, 스포츠, 재난, 유통, 금융 등 유망업종에서 빅데이터 선도 프로젝트를 추진(3년간 350억원, 민간 매칭)

1) 미래창조과학부, 빅데이터 기반 미래예측 및 전략수립 지원계획, 2014.

2) 미래창조과학부, 창조경제 실현을 위한 K-ICT 전략, 2015.

- 빅데이터로 시민체감형 도시문제(교통, 주차 등) 해결 시범사업 추진(17년까지 6곳)
- 빅데이터 산업 활성화와 개인정보 보호가 조화되도록 개인정보보호법 등 관련 법제 개정 추진
- 우리나라는 K-ICT 전략을 기반으로 2019년까지 글로벌 빅데이터 3대 강국 진입을 목표로 하고 있음
- 이외에도 정부통합전산센터 빅데이터분석과, 국토교통부 국토정보정책과, 한국정보화진흥원 빅데이터전략센터 등에서 빅데이터 관련 사업을 추진하고 있음

### 3. 지역별 빅데이터 추진 사례

#### 1) 서울특별시

- 서울시는 타 지자체에 비해 여러 측면에서 압도적으로 많은 데이터가 생성되고 있는데, 이러한 데이터 중 접근이 용이한 교통 부문에 중점을 두어 빅데이터를 활용하였으며 이를 통해 효율성을 제고함
- 교통부문(예: 심야버스 등) 중에서도 가시적인 효과를 얻고, 단계적으로 확대해 나가는 전략으로 추진함<sup>3)</sup>
  - 2013년 : 데이터 중심의 공유 및 분석 기반 마련을 위한 정책 과제 수요조사, 중장기 추진 전략수립 및 시범과제 추진
  - 2014년 : 추진사업 대상 데이터 수집, 공유 및 분석, 전담조직구성 및 내부역량 강화와 제도·지침 마련
  - 2015년 : 예측기반의 시정운영 혁신 및 미래 복합행정서비스 구현을 위해 데이터 공유·분석체계를 확대하고, 다차원적이고 지능형의 빅데이터 활용체계 구축
- 또한 서울시는 빅데이터 관련 사업의 효율적인 추진을 위해 정보기획관 아

3) 경남신문, 빅데이터를 활용한 지방행정, 2015년 7월 23일자.



래 전담부서를 조직했는데 4개팀(빅데이터 전략팀, 데이터개발팀, 통계정보팀, 통계조사팀) 30여명으로 구성되어 있고, 교통, 산업경제, 도시관리, 안전, 문화관광·플랫폼 고도화 사업 등을 추진하고 있음<sup>4)</sup>

- 아울러 서울시는 서울시인재개발원에 빅데이터 관련 과목을 편성해 빅데이터 교육을 전 부서로 확대하고 있는 상황임
- 이외에도 서울시는 신한카드와 업무협약을 통해 카드 소비 데이터를 활용한 매출정보를 제공받아 유동인구, 교통정보, 휴·폐업 정보 등 골목상권의 최신 데이터를 세밀하게 분석하여 다양한 정책개발에 활용할 계획임
- 이러한 빅데이터 활용 정책을 제도적으로 뒷받침하기 위해 2015년 4월 다음과 같은 주요 내용을 담은 관련 조례안을 발의했으나 계류 중에 있음
  - 빅데이터 활용을 위한 기본계획 수립 및 시행 의무화
  - 서울시 빅데이터 위원회 구성
  - 빅데이터 수집·분석, 활용 사무를 총괄하는 빅데이터 책임관제 도입
  - 빅데이터센터 설치 및 빅데이터 전문인력 양성 등

## 2) 경기도

- 경기도는 빅데이터와 관련하여 민선 6기 공약사항으로 빅파이(Big-Fi) 프로젝트를 추진하고 있는데 빅파이 프로젝트는 4개 분야 12개 추진과제 및 29개 세부사업으로 구성되어 있음<sup>5)</sup>
  - 4개 분야 : 도정 혁신, 투명한 도정, 민간활용 지원, 공통기반 구축
  - 12개 추진과제 : 일하는 방식 개선, Best Practice 시범사업 추진, 의사결정 지원 플랫폼 구축, 빅파이 가치 공감대 확산, 데이터 개발 플랫폼 구축, 데이터 개발 사업화 지원, 비즈니스 모델 발굴, 전문인력 양성, 데이터 생태계 조성, 빅데이터

4) 서울특별시청 홈페이지(www.seoul.go.kr).

5) 경기콘텐츠진흥원(빅파이 추진단), 빅파이추진단 중·장기 경영목표, 2015.1.28.

활용 문화 확산, 개인정보 보호 방안 마련, 데이터 활용 제도 개선

- 경기도는 빅데이터 사업 추진을 위해 빅데이터담당관을 두고 담당관을 포함해 16명의 인력이 관련 사업을 추진하고 있음<sup>6)</sup>
- 또한 빅데이터 관련 사업의 효과적인 추진을 위해 2015년 3월 5개팀으로 구성된 빅파이 추진단을 조직했고, 시범사업 발굴, 전문인력 양성, 빅데이터 활용문화 확산을 위한 활동을 추진하고 있음
- 이외에도 경기도는 빅데이터 활용을 통한 행정의 효율성 및 지역경제 발전을 도모를 목적으로 하는 '빅데이터 활용에 관한 조례'를 2015년 3월 제정함
  - 경기도 관련 공공기관의 빅데이터 활용을 통한 지역발전에 기여, 빅데이터 책임관 지정, 빅데이터 활용에 관한 기본계획 의무화, 위원회 구성, 빅데이터 활용기반 구축, 실태조사 등이 주요 내용임

### 3) 부산광역시 해운대구

- 부산시 해운대구청은 다양해진 정보를 효과적으로 활용하기 위해 2013년 1월 기획조정실 내에 3명으로 구성된 빅데이터 분석팀을 신설함<sup>7)</sup>
- 빅데이터 분석팀에서는 해운대 관광객의 소셜데이터 분석을 통해 과학적 관광정책을 마련한 바 있음<sup>8)</sup>
  - 2013년 : 5월 ~ 10월까지 트위터와 페이스북, 블로그 등 SNS에 해운대라는 키워드 분석을 통해 해운대 관광의 매력요인(해운대 해수욕장, 달맞이길 등)과 방해요인(비싼 숙박요금, 교통 체증, 복잡한 버스 노선 등) 분석 ⇒ 부당요금 근절 및 게스트하우스 확충 시책
  - 2014년 : 5월 ~ 10월까지 트위터와 페이스북, 블로그 등 SNS 재분석을 통해 관광서비스 개선 및 해운대 야경 7선 선정

6) 경기도청 홈페이지(www.gg.go.kr).

7) 부산광역시 해운대구청 홈페이지(www.haeundae.go.kr).

8) 경남신문, 빅데이터를 활용한 지방행정, 2015년 7월 21일자.

- 이외에도 수년간 쌓여온 행정 데이터 활용을 통해 분야별, 부서별 민원의 주요 키워드 및 트렌드, 연관어를 분석하고 연도별 변화 추이 등 통계 상황과 특성을 파악하여 예상 가능한 민원을 사전에 대처할 수 있도록 활용함
  - 해운대의 경우 해수욕장 및 관광지가 몰려 있어 쓰레기 무단투기에 대한 문제가 심각한데 빅데이터 분석을 통해 지역별 맞춤형 정책을 시행함
  - 지역 구직 희망자 데이터 분석, 주·정차 위반사례를 분석하는 등 빅데이터 분석을 정책 발굴의 기초 자료로 활용하기 위한 노력하고 있음
- 향후 교통, 복지, 문화 등 각종 데이터 분석 결과를 정책에 활용할 계획에 있음

#### 4) 창원시

- 창원시는 2014년 6월 빅데이터를 이용한 업무혁신과 아이디어 활용 정책을 제안해 정부 3.0 선도과제로 선정된 바 있고, 진해 군항제 관광객 관련 빅데이터 분석을 통해 유의미한 결과를 도출한 바 있음<sup>9)</sup>
- 이러한 경험을 바탕으로 2015년에는 창원시 ICT 융합 및 스마트 환경을 구축하기 위해 정보통신담당관이 주축이 되어 빅데이터 플랫폼 구축 및 분야별 모델 개발과 분석 사업을 추진하고 있음<sup>10)</sup>
  - 빅데이터 플랫폼 구축 : 데이터 마트 구성 및 빅데이터 기반의 분석 플랫폼 설계/구현
  - 관광서비스 모델 구축 : 분석을 통한 관광권역 설정 및 관광정책 발굴
  - CCTV 사각지대 분석 : 창원시 행정데이터 및 외부데이터 교차분석
  - 시내버스 노선분석 : 기존버스 노선 조정 정책 결정
  - 전통시장 분석 : 이용패턴 분석을 통한 배후 수요지 분석
- 창원시는 향후 다양한 현안문제를 파악하고, 이를 활용하기 위해 빅데이터를 활용할 계획을 가지고 있음

9) 경남신문, 창원시 빅데이터로 정책 효율성 높인다, 2015년 5월 21일자.

10) 창원시청 홈페이지([www.changwon.go.kr](http://www.changwon.go.kr)).

## 4. 국외 빅데이터 추진 사례

### 1) 미국 로스앤젤레스 보육 서비스 관련 사기행위 근절

- LA 카운티는 미국 내 카운티 중에서 가장 많은 인구가 거주하고 있고, 캘리포니아 총 주민 중 약 27%정도의 비중을 차지하고 있고, 예산은 약 230억 달러로 사회, 의료 관련 및 법 집행 서비스를 위해 지출됨
- LA 카운티는 육아와 관련된 보육서비스에서 사기 범죄가 급격히 증가하는 것을 우려하여 빅데이터 분석을 통해 보조금 수령자 및 서비스 공급자 등을 파악하여 부정행위와 연관성이 있는 대상 파악, 기금 손실 우려 분야 예측 등의 빅데이터 분석을 실시함<sup>11)</sup>
- 보육서비스와 연계된 주요 부정행위는 유령직원 신고, 허위 프로그램 추진에 따른 부정 수급 범죄임
  - 실제 근무하지는 않지만 근무하는 것처럼 위장하여 직원들을 신고하고, 이에 따른 급여를 부정 수급함
  - 존재하지 않는 유령 회사를 만들고, 이 유령회사에서 보육 프로그램을 추진하는 것처럼 위장하고, 수혜자들과 공모해 복지예산을 부정수급하는 등의 범죄를 일으킴
- 정부기관은 방대한 자료를 기초로 보육서비스와 관련된 이상 징후를 찾고, 과거 정보를 기반으로 예측할 수 있는 모델을 개발하여 부정수급 패턴과 유사 범죄 상황을 파악함
- 세부적으로는 주소, 전화번호, 계좌번호, 차량 정보, 영유아 보육센터 관련 데이터간 연관관계를 파악하고, 개발한 모델을 통해 관련 고용데이터, 보육시설 사업허가 데이터 등을 분석함
- 또한 DB와 DW 및 GIS 정보를 함께 분석하여 효과적인 분석을 실시함
  - 소셜데이터 및 GIS 정보를 통합하여 영유아 위치와 보육시설의 위치를 지도상에 표시함

11) 노규성, 사례분석을 통한 지방행정의 빅데이터 활용 전략, 2014.12 p.92.

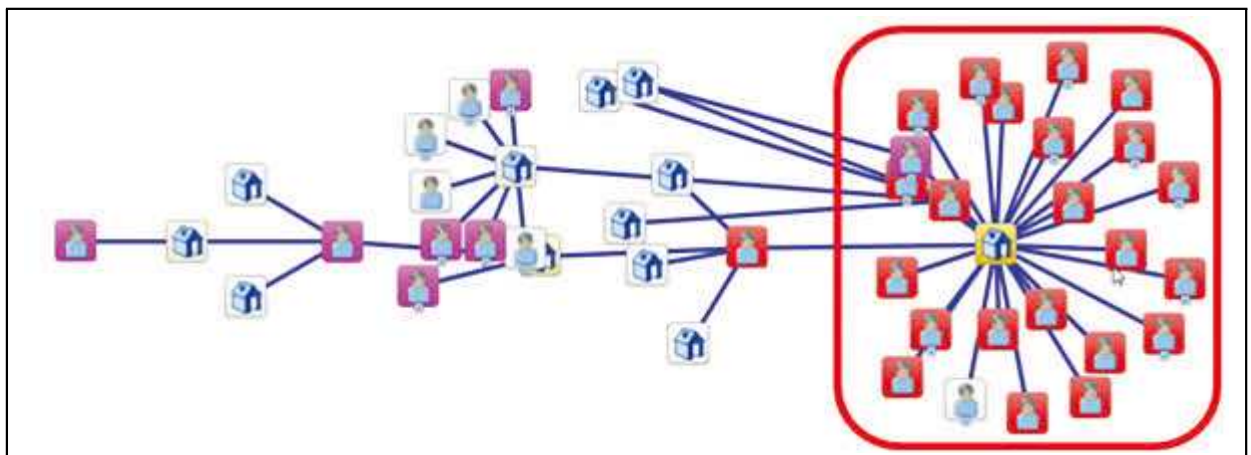
- 지리적으로 보육서비스 제공이 불가능한 거리에 위치해 있는 유아와 보육시설이 존재하는 것을 파악함



자료)한국정보화진흥원, 더 나은 미래를 위한 데이터 분석, 2013.4.

(그림2) 보육시설 관련 부정수급 사례\_영유아 위치 vs. 보육시설

- 또한 보호자의 명의를 다른데 주소나 전화번호가 동일한 영유아가 있는 것이 파악되면서 조직적인 범죄인지 여부를 조사하여 복지서비스 이용 영유아 중복 및 보육기관과 보호자가 공모하여 복지에산을 부정 수급한 범죄사실을 밝혀냄



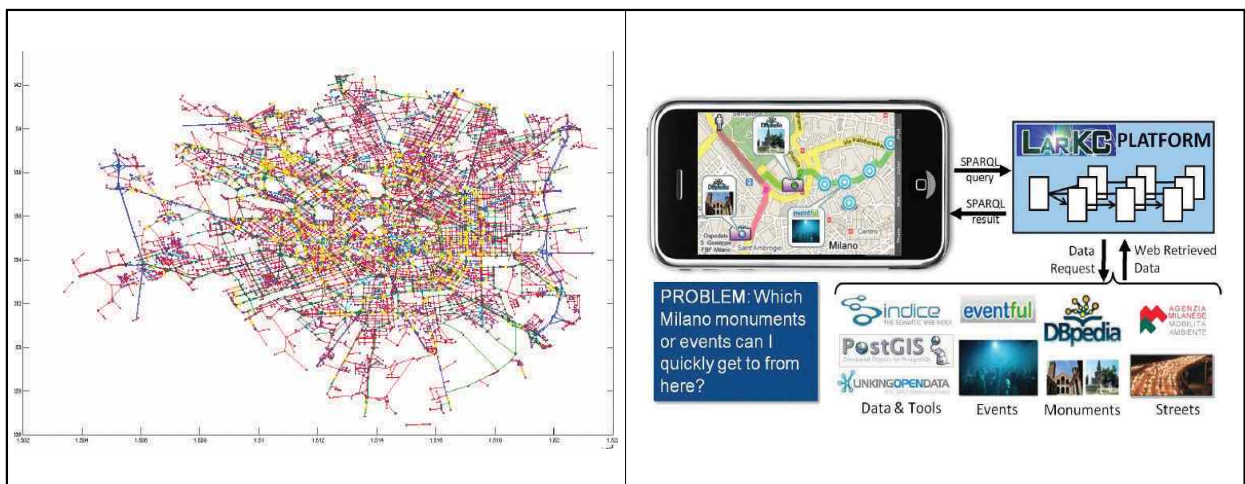
자료)한국정보화진흥원, 더 나은 미래를 위한 데이터 분석, 2013.4.

(그림3) 보육시설 관련 부정수급 사례\_대상 영유아 vs. 주소/전화번호 등

- 이러한 노력을 통해 영유아 복지관련 사기 집단을 효과적으로 적발(85%)할 수 있게 됨

## 2) 밀라노 교통상황 예측

- 밀라노에서는 민·관·산이 협력하여 차량의 비효율적인 이동을 막고, 이를 통해 에너지 낭비 저감 및 자동차 매연 감소를 실현하고자 빅데이터 분석을 통한 교통서비스를 개발함
- 밀라노에서는 교통상황 예측을 위해 2007년 5월부터 2009년 7월까지 경찰서의 교통사고 정보, 기상청 날씨 데이터, 도로교통 흐름, 도로공사 및 주변 건축물 데이터, 기타 행사 등의 데이터를 수집함
- 교통 흐름양 및 속도 파악을 위해 5분 간격으로 도로상황 정보를 측정하여 데이터화 했으며, 데이터 분석을 위해 도로 900여 곳에 센서를 부착함
- 도로 곳곳에 부착된 센서는 2~24시간 도로 교통상황 예측이 가능한데 밀라노 도시 전체 교통흐름 정보 분석에는 불과 0.1초가 소요됨
- 또한 예외적으로 발생하는 정보에 대해서는 실시간으로 데이터화하여 시스템에 반영될 수 있도록 함



자료) 한국정보화진흥원, 빅데이터로 진화하는 세상, 2012.5.

(그림4) 밀라노 시내 교통센서 지도(좌), 교통생활 애플리케이션(우)

- 교통의 흐름에 영향을 끼치는 다양한 요소들인 시간, 기후 등을 고려하여 효율적인 길찾기 서비스를 제공함
- 또한 설치되어 있는 센서를 통해 획득되는 데이터를 통해 예외상황(사고, 공사 등)을 반영한 서비스 제공이 가능함
- 이를 기반으로 향후 차량 속도 데이터와 거리 데이터를 실시간으로 활용하여 보다 정확한 교통 흐름을 제공하기 위해 노력하고 있음

### 3) 일본 영상정보 활용을 통한 현지 상황 파악

- 일본에서는 국토교통성이 주축이 되어 지진 관측망과 전국 15,000대의 CCTV 영상 정보를 활용해 현지 상황 파악과 구조와 관련된 신속한 활동을 목적으로 하는 시스템을 2007년 구축함
- 전국 약 700개소인 지진 관측망에서 진도 4 이상의 지진이 관측되면 국토교통성의 IP CCTV 네트워크 시스템을 통해 해당 지역 CCTV 카메라를 표시하는 아이콘의 색깔이 변하게 됨
- 진도 4 이상의 지진이 관측된 경우 CCTV 정보 검색, 현지 도로 상황 파악, 여진 발생 여부 등에 따라 대기 또는 재해대책본부 등과 신속하게 연계될 수 있도록 함



자료) 한국정보화진흥원, 빅데이터로 진화하는 세상, 2012.5.

(그림5) CCTV 카메라를 통한 도로 상황 파악 시스템

- 이러한 현황 정보는 위치, 지진강도, 피해상황, 여진 발생 여부 등으로 구분하여 저장되고, 추후 지진 발생을 예측하는 데에도 활용함
- 2009년에는 CCTV 카메라의 활용 고도화를 위해 게릴라성 호우로 인해 하천 수위가 급격히 상승하여 침수 피해가 발생하는 상황에 효과적으로 대처하기 위해 피해 시 포착된 영상을 활용할 수 있도록 하고, 기록된 정보는 향후 예측에도 활용하여 피해를 최소화하는데 기여할 수 있도록 하고 있음



## Ⅲ

## 경상남도 빅데이터 사업 및 체계 구축 방향

## 1. 경상남도 빅데이터 환경

- 경상남도는 정부 3.0의 기조(국민 개개인 중심, 확장된 민주성, 개방공유·협력·소통, 양방향·맞춤형 제공, 무선 인터넷 스마트 모바일)와 함께 하기 위해 경상남도가 보유하고 있는 공공데이터를 우선 개방함
- 경상남도 공공데이터는 '경상남도 공공데이터 개방 포털'을 통해 제공되고 있는데 총 18개 분야(복지, 관광, 행정, 산업 등)의 데이터를 제공하고 있고, 교통, 도시/주택, 보건, 환경 등 다양한 분야에 걸쳐 공공데이터 개방을 위해 노력하고 있음
- 나아가 경상남도는 정부 3.0 데이터 기반의 과학적 정책 수립에 활용할 수 있는 빅데이터 기술 도입을 적극적으로 검토하고 있음

## 2. 경상남도 빅데이터 사업 추진 방향 설정을 위한 의견조사

## 1) 조사개요

- 경상남도의 경우 빅데이터 사업에 관한 추진 의사는 있으나, 현재까지 추진한 경험이 전무하므로 관련분야 실무전문가를 대상으로 빅데이터 사업 추진 방향 설정 및 추진 분야 파악을 위한 의견을 수렴함
- 본 조사는 장기적 차원에서의 조사이므로 각 분야별(예 : 복지, 산업, 교통, 생활 등) 실무자에 대한 조사는 배제하고, 빅데이터의 필요성에 대한 인식과 관심이 있으면서 경상남도 전반의 분야에 경험이 있는 경상남도 정보통계담당관실 실무직원과 시·군의 정보통신 업무를 담당하고 있는 실무직원을 대상으로 조사를 실시함
- 본 조사는 개별 면접조사를 실시하여 조사 목적의 명확성 전달 및 회수율을

높일 수 있도록 함

(표1) 경상남도 빅데이터 사업 추진 방향 설정의 위한 의견 조사 개요

| 구분      | 내용  |
|---------|---|
| 조사대상    | ○경상남도 정보통계담당관실 및 18개 시·군 정보통신 관련 공무원  |
| 조사시기    | ○2015년 11월 25일 : Pilot Test<br>○2015년 12월 2일 ~ 8일 : 본 조사                                      |
| 배포부수    | ○38부  |
| 회수부수    | ○35부(회수율 92.1%)   |
| 유효부수    | ○30부  |
| 조사방법    | ○면접조사 및 설문조사  |
| 주요 조사내용 | ○경상남도 빅데이터 사업 추진 방향에 관한 우선순위<br>○사업 추진 방향별 초점을 맞추어야 할 사항<br>○경상남도 빅데이터 시범사업 추진 시 우선 분야 및 세부사업 |

- 효과적인 조사를 위해 경상남도 현안, 도지사 공약, 이슈 사항 등을 고려하여 경상남도 빅데이터 추진 방향(안)과 사업 추진분야(안)를 마련하고, 이를 기반으로 조사를 실시함
- 경상남도 빅데이터 추진 방향(안)은 크게 기타를 제외한 8개 부문으로 구분하고 각 방향별 주요 세부사업을 제시하여 향후 경상남도 빅데이터 사업은 어떠한 방향에 중점을 두는 것이 효과적인지를 조사함

(표2) 경상남도 빅데이터 사업 추진 방향(안)

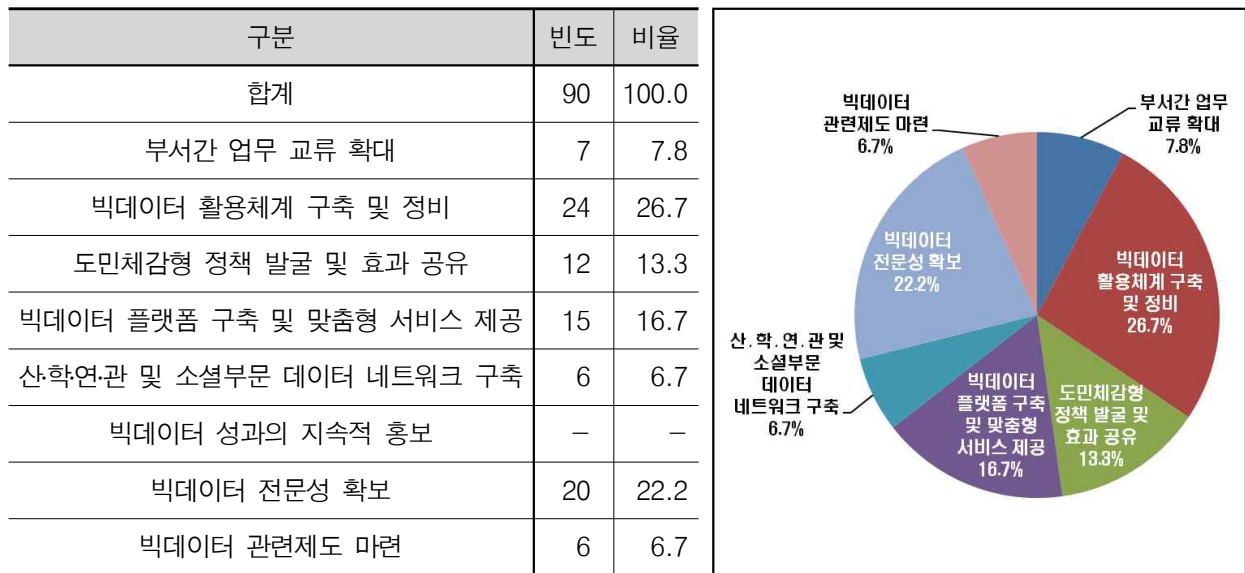
| 사업 추진 방향                 | 세부방향  |
|--------------------------|---|
| 부서간 업무교류 확대              | <ul style="list-style-type: none"> <li>○ 실무협의회 구성 및 운영</li> <li>○ 도 및 18개 시·군 데이터 공유 체계 구축</li> <li>○ 도 및 시·군 현안문제 공유체계 구축</li> </ul>                                       |
| 빅데이터 활용체계 구축 및 정비        | <ul style="list-style-type: none"> <li>○ 빅데이터 분석을 고려한 데이터 축적</li> <li>○ 문제 해결을 위한 데이터 연계 체계 구축</li> <li>○ 분석서비스 제공 모델 개발</li> </ul>   |
| 도민체감형 정책 발굴 및 효과 공유      | <ul style="list-style-type: none"> <li>○ 빅데이터 분석을 통한 현안 사업 발굴</li> <li>○ 도민체감 정책 관련 아이디어 뱅크 구축</li> <li>○ 빅데이터 관련 교육/세미나/추진 결과 공유</li> </ul>                                |
| 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공 | <ul style="list-style-type: none"> <li>○ 빅데이터 개발 플랫폼 구축</li> <li>○ 빅데이터 개방조정 협의회 구성</li> <li>○ 수요도에 따른 빅데이터 개방 확장</li> <li>○ 산업별 관련 공공데이터 세분화</li> </ul>                    |
| 산학연관 및 소셜부문 데이터 네트워크 구축  | <ul style="list-style-type: none"> <li>○ 비정형데이터 DB화</li> <li>○ 공공SNS 데이터마트 구축</li> <li>○ 산·학·연·관 데이터 네트워크 구축</li> <li>○ 일자리 창출, 중소기업 지원, 창원지원 등을 위한 데이터서비스 체계 구축</li> </ul> |
| 빅데이터 성과의 지속적인 홍보         | <ul style="list-style-type: none"> <li>○ 빅데이터 성과 및 사례 발표대회 개최</li> <li>○ 빅데이터 아이디어 공모전 개최</li> <li>○ 언론매체 홍보</li> </ul>   |
| 빅데이터 전문성 양성              | <ul style="list-style-type: none"> <li>○ 분야별 빅데이터 교육과정 개설</li> <li>○ 산·학 연계를 통한 취업 지원</li> </ul>  |
| 빅데이터 관련제도 마련             | <ul style="list-style-type: none"> <li>○ 개인정보 비식별화 및 암호화 방안 마련</li> <li>○ 개인정보 보호를 고려한 데이터 수집 체계 정비</li> <li>○ 공공데이터 이용 관련 조례 제정</li> <li>○ 공공데이터 제공을 위한 기준 마련</li> </ul>   |
| 기타                       | <ul style="list-style-type: none"> <li>○ 빅데이터 센터 운영</li> <li>○ 위원회 및 자문단 구성</li> </ul>  |

○ 이와 함께 빅데이터 사업 추진 시 우선적으로 추진해야할 분야(안전, 복지, 산업, 관광 등) 및 추진 사업에 대한 조사도 병행하여 효과적인 경상남도 빅데이터 사업 추진 계획 수립에 기반을 마련함

## 2) 조사결과

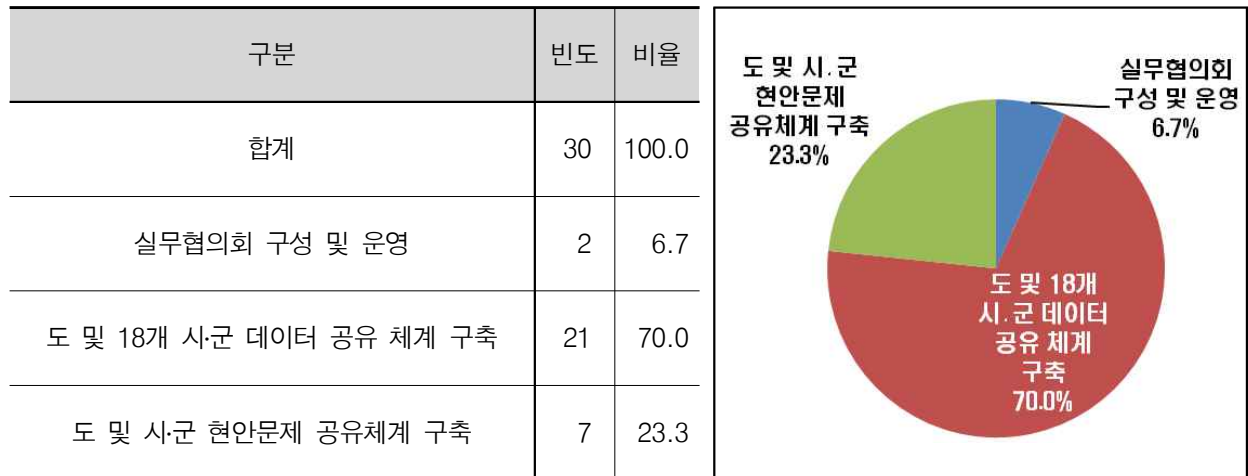
- 경상남도 빅데이터 사업 추진 방향 설정 및 추진 관련 설문 조사 내용 중 빅데이터 추진 방향과 관련해서는 빅데이터 활용체계 구축 및 정비(26.7%), 빅데이터 전문성 확보(22.2%), 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공(16.7%)이 가장 중요하다고 응답함

(표3) 빅데이터 사업 추진 방향



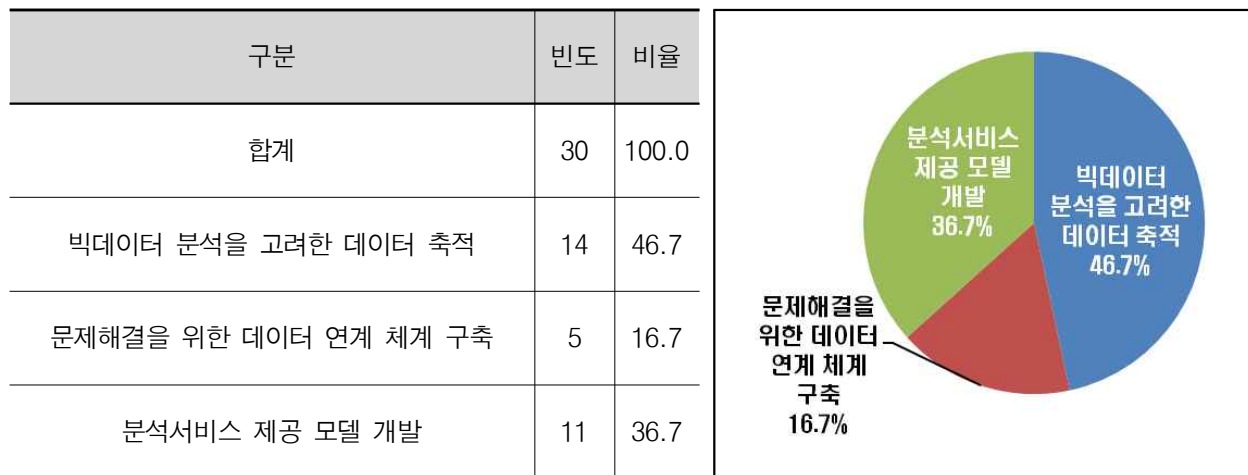
- 이러한 결과를 볼 때 빅데이터 사업과 관련해서는 축적되어 있는 데이터의 정비와 전문인력의 보유, 도민들이 체감할 수 있는 서비스를 제공하는데 중점을 두어야 하는 것으로 파악됨
- 빅데이터 사업 추진 방향 중 부서간 업무 교류 확대와 관련해서는 도 및 18개 시·군간 데이터의 공유를 위한 체계를 구축하는 것이 중요(70.0%)하다는 의견이 대부분을 차지하는 것으로 분석됨

(표4) 부서간 업무 교류확대



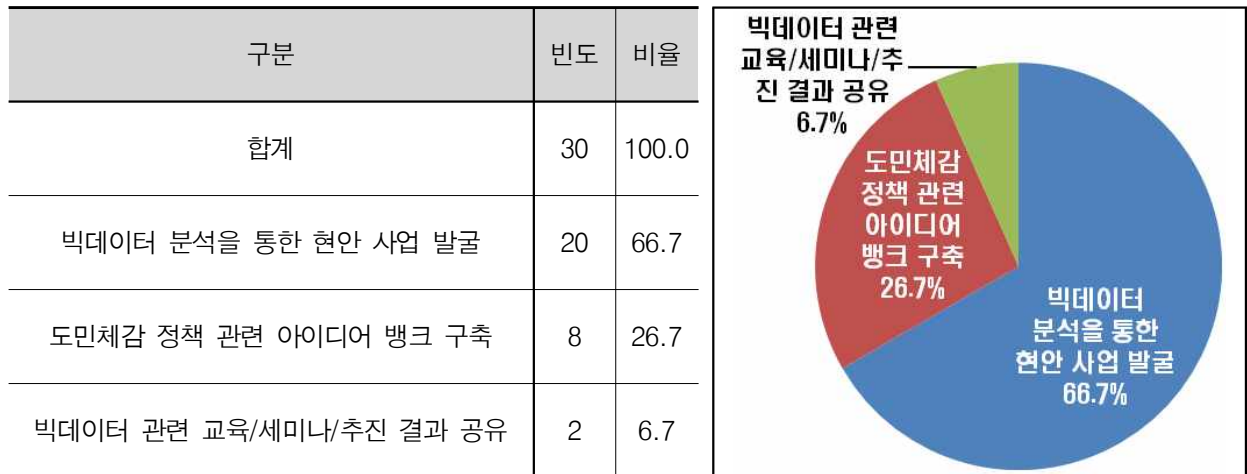
○ 빅데이터 활용체계 구축 및 정비부문의 세부 방향에 있어서는 빅데이터 분석을 고려한 데이터 축적이 가장 중요하다는 의견(46.7%)이었고, 분석서비스 제공 모델 개발도 중요하다고 생각하는 의견(36.7)이 비교적 많은 것으로 조사됨

(표5) 빅데이터 활용체계 구축 및 정비



○ 도민체감형 정책 발굴 및 효과 공유와 관련해서는 빅데이터 분석을 통한 현안 사업 발굴이 가장 중요(66.7%)한 것으로 조사됨

(표6) 도민체감형 정책 발굴 및 효과 공유



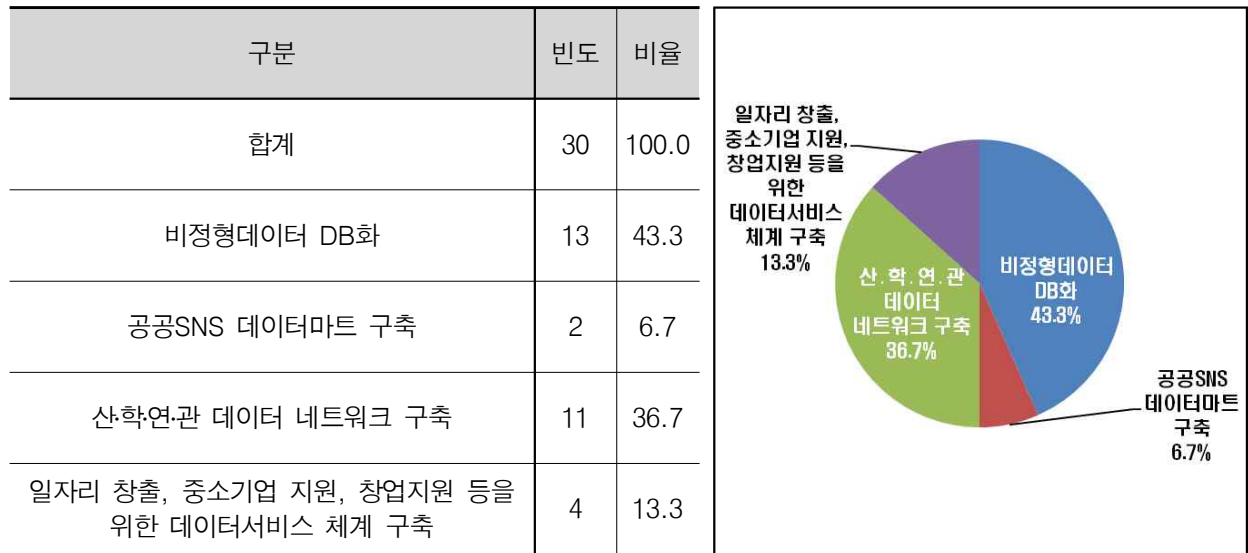
○ 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공 부문에 있어서는 서비스 제공의 기반이 되는 빅데이터 플랫폼 구축이 가장 많은 비중(63.3%)을 차지했고, 다음으로 수요도에 따른 빅데이터 개방 확장, 산업별 관련 빅데이터 세분화 등의 순으로 중요도가 높은 것으로 조사됨

(표7) 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공



○ 산·학·연·관 및 소셜부문 데이터 네트워크 구축 부문에서는 비정형데이터의 DB화의 비중(43.3%)이 가장 높은 것으로 조사되었고, 다음으로 산·학·연·관 데이터 네트워크 구축, 일자리 창출·중소기업 지원·창업지원 등을 위한 데이터서비스 체계 구축의 순으로 중요하다는 의견임

(표8) 산·학·연·관 및 소셜부문 데이터 네트워크 구축



○ 빅데이터 성과의 지속적 홍보 부문에서는 빅데이터 성과 및 사례 발표대회를 개최하는 것이 가장 중요(50.0%) 하다는 의견이었고, 아이디어 공모전 개최도 적지 않은 비중(36.7%)을 차지하는 것으로 분석됨

(표9) 빅데이터 성과의 지속적 홍보



○ 전문성 확보 부문에 있어서는 전담조직의 구성과 외부전문가 및 전문기관과의 교류체계 구축이 비슷하게 조사되었는데 전문성 확보의 중요성을 보여주는 조사 결과라 판단됨

(표10) 빅데이터 전문성 확보



- 빅데이터 관련제도 마련과 관련해서는 국가적으로도 고민하고 있는 개인정보 보호를 고려한 데이터 수집 체계 정비 부문이 중요하다는 의견(53.3%)이 가장 많았음
- 개인정보 보호와 관련해서는 사업 추진 시에도 많은 제약으로 작용할 수 있으므로 지속적인 검토가 필요한 부분이라 판단됨

(표11) 빅데이터 관련제도 마련



- 사업추진 방향 설정과 관련하여 종합하면 빅데이터 사업 추진을 위해서는 기반을 조성하는 측면(활용체계 및 플랫폼 구축, 전문성 확보)에 초점을 맞추는 것이 중요한 것으로 판단됨
- 또한 현재까지는 시범사업도 추진하지 않은 상태이므로 데이터 네트워크 구축이나 성과 홍보에 대해서는 중요도가 많이 낮은 것으로 분석됨



- 따라서 경상남도 빅데이터 사업 추진 방향 설정 시 이러한 조사 내용을 충분히 반영하여 설정함

### 3. 경상남도 빅데이터 사업 추진 방향 설정

#### 1) 비전 및 목표

- 경상남도는 빅데이터와 관련한 경험이 부족하므로 초기 단계에 적합한 목표로 빅데이터 활용체계 구축 및 정비, 빅데이터 전문성 확보, 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공으로 설정함
- 이러한 목표 달성을 통해 「Big Data 기반 도민 삶의 질 개선」이라는 비전을 실현시킬 수 있을 것으로 판단되고, 나아가 안전하고, 깨끗하며, 행복한 경남을 만들어 나가는데 기반이 될 수 있을 것이라 판단됨



(그림6) 경상남도 빅데이터 사업 추진 방향(비전/목표/전략방향)

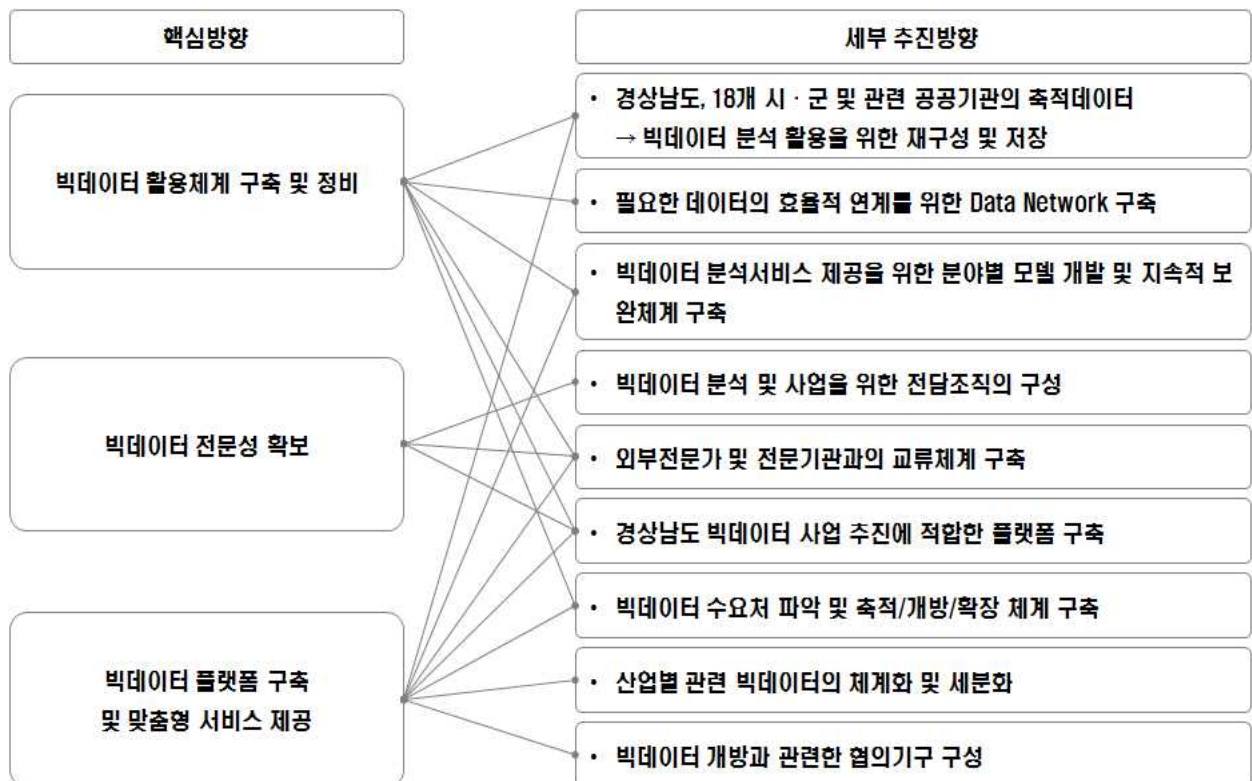
- 빅데이터 활용체계 구축 및 정비 부문에 있어서의 추진 전략 방향은 기존에 무분별하게 축적되던 데이터들을 빅데이터 분석을 고려하여 수집·저장하고,

이를 기반으로 효과적인 서비스 제공을 위한 모델 개발, 데이터 연계 체계 구축 등으로 전략방향을 설정함

- 빅데이터 전문성 확보 부문에 있어서는 전담조직 구성, 외부전문가 및 전문기관과의 교류 체계 구축으로 전략방향을 설정함
- 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공을 위해서는 빅데이터 개발 플랫폼 구축, 수요도에 따른 빅데이터의 개방 및 확장, 산업별 관련 빅데이터 세분화, 빅데이터 개발 및 조정협의회 구성으로 전략 방향을 설정함

## 2) 빅데이터 사업 추진 방향

- 경상남도 빅데이터 사업 추진은 성과보다 기반을 조성하는데 초점을 맞추어야 할 것으로 판단됨
- 따라서 활용체계 구축, 전문성 확보, 플랫폼 구축 및 맞춤형 서비스 제공을 위한 세부 추진 방향을 설정할 수 있음



(그림7) 경상남도 빅데이터 사업 추진 방향

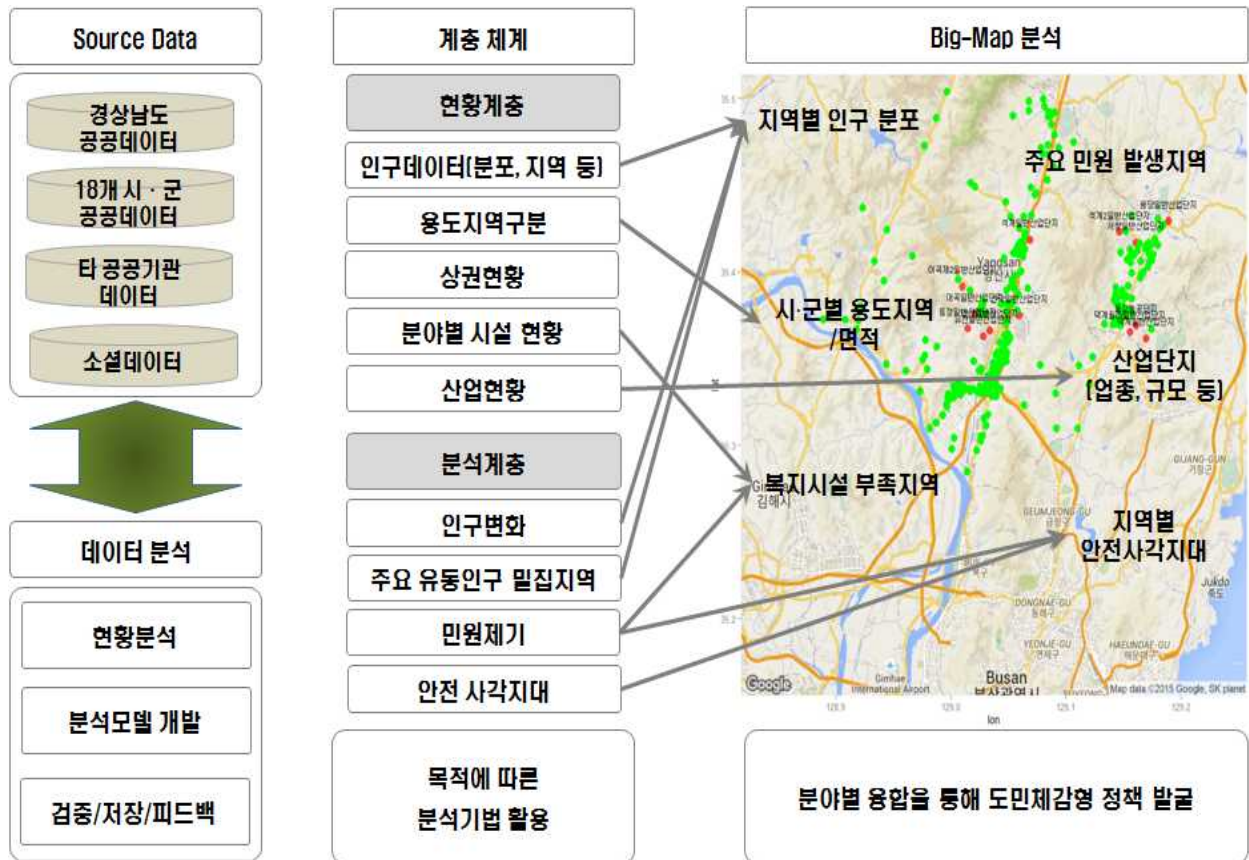
- 빅데이터 활용체계 구축 및 정비와 관련해서는 다음과 같은 세부 추진 방향을 설정할 수 있음
  - 경상남도, 18개 시·군 및 관련 공공기관에서 기 축적되어 있는 데이터를 빅데이터 분석에 활용할 수 있도록 재구성한 데이터 축적
  - 재구성한 분야별 데이터가 필요 시 효과적으로 연계될 수 있는 Data Network 구축
  - 빅데이터 분석서비스 제공을 위한 분야별 모델 개발 및 지속적 보완 체계 구축
- 빅데이터 전문성 확보를 위해서는 다음과 같은 세부 추진 방향을 설정할 수 있음
  - 빅데이터 분석 및 사업을 위한 전담조직의 구성
  - 정기적으로 외부전문가의 자문을 구할 수 있는 체계 구축
  - 빅데이터 관련 외부전문기관(예: 한국정보화진흥원, 경기도 빅파이 추진단 등)과의 교류체계 마련
- 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공을 위해서는 다음과 같은 세부 추진 방향을 설정할 수 있음
  - 경상남도 빅데이터 사업 추진에 적합한 플랫폼 구축
  - 빅데이터 분석을 요구하는 수요처를 파악하고, 이를 위한 빅데이터의 축적·개방·확장할 수 있는 체계 마련
  - 각 산업별 관련되는 빅데이터의 체계화 및 세분화
  - 빅데이터 개방과 관련한 협의 기구 구성
- 종합하면 경상남도 빅데이터 사업 추진 방향은 분야별 세분화된 방향을 제시하고 있으나 세부 방향은 상호 연계될 수 있으므로 이를 고려하여 사업을 추진하는 것이 타당할 것으로 판단됨

### 3) 경상남도 빅데이터 분석 및 서비스 계층 구성 방안

- 빅데이터 분석 및 서비스 모델을 구축하기 위해서는 기초데이터라고 할 수 있는 경상남도, 18개 시·군 및 관련 공공기관의 데이터 및 분야별 소셜데이

터의 수집이 기반이 되어야 함

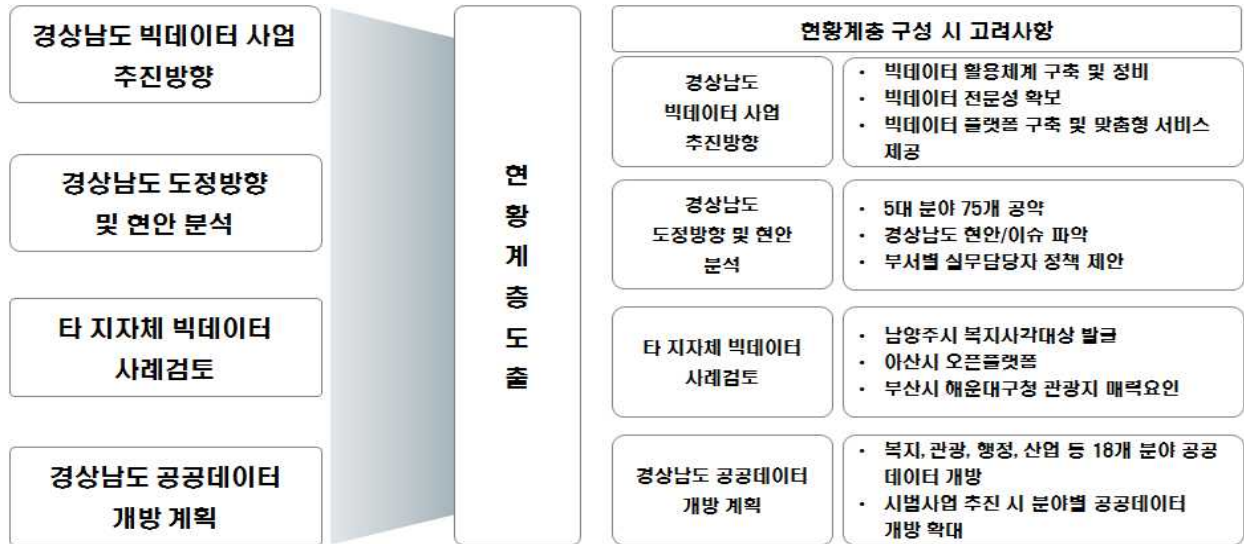
- 기초데이터를 기반으로 현황을 분석하고, 분야별 분석을 위한 모델을 개발하며 모델을 통해 제시되는 결과들에 대한 검증, 저장, 피드백 등의 과정을 거치게 됨
- 즉, 크게 현황계층과 분석계층으로 구분하고 각 계층간 연계 분석을 통해 효과적인 결과를 제시하고, 이를 기반으로 도민들이 체감할 수 있는 정책을 발굴할 수 있도록 함



(그림8) 경상남도 빅데이터 사업 추진 계층 구성

#### 4) 경상남도 빅데이터 사업 추진을 위한 계층 구성 방안

- 경상남도 빅데이터 사업의 효율적인 추진을 위해서는 데이터 관점에서 개념적인 계층(Layer)의 구성이 필요한데 앞서 설명한바와 같이 현황계층과 분석계층으로 분류할 수 있음



(그림9) 빅데이터 사업 추진을 위한 현황 계층

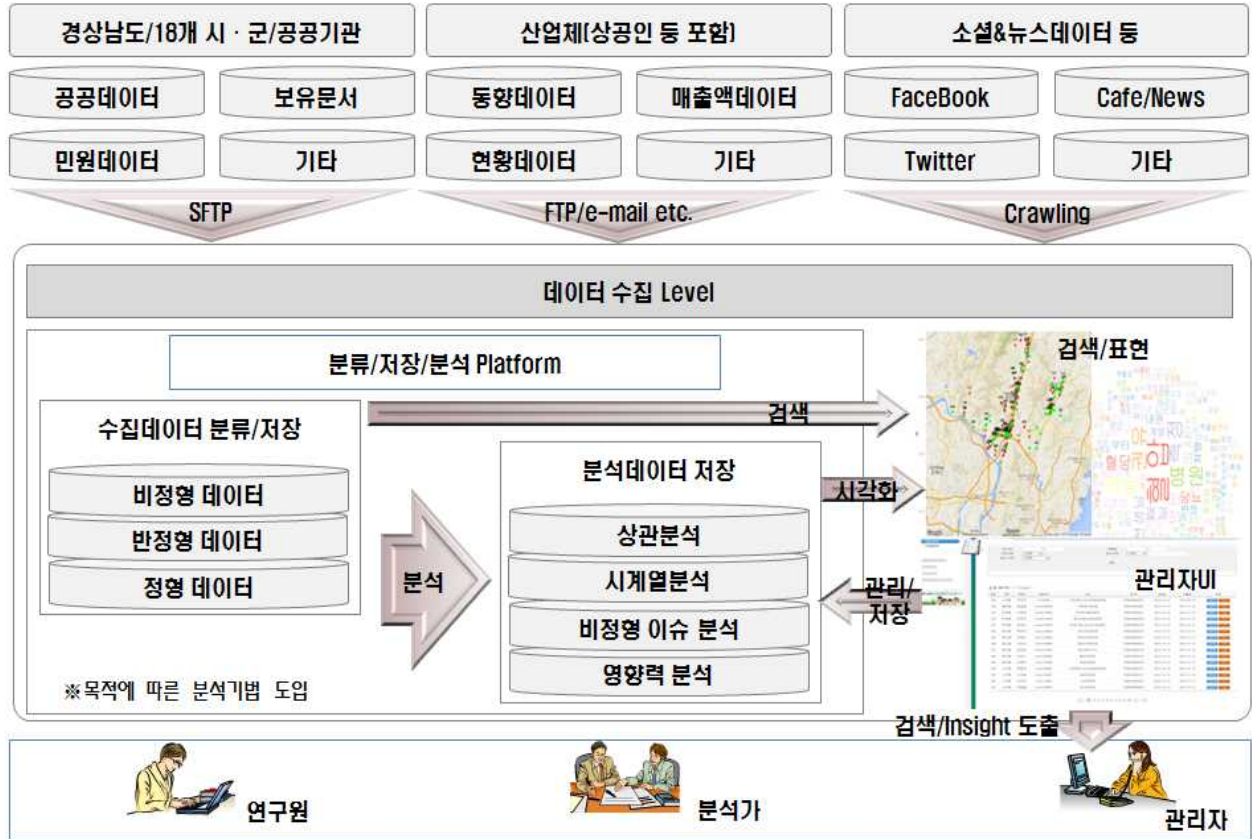
- 먼저 경상남도 빅데이터 사업 추진의 기반이 되는 현황계층은 경상남도 빅데이터 사업 추진 방향, 경상남도 도정방향 및 현안, 타 지자체 빅데이터 사례, 경상남도 공공데이터 개방 등의 차원에서 도출할 수 있음
- 현황계층에서는 기초데이터의 수집, 저장, 기존 사업 모델 검토, 경상남도 빅데이터 사업 추진 시 요구되는 데이터 정의 및 수집 방법 등을 전제로 구성됨
- 다음으로 분석계층에서는 기 확보된 기초데이터, 기존 사업 모델 등을 기반으로 문제해결을 위한 분석기법 적용 및 모델을 개발하는 단계로 볼 수 있음
- 분석계층에서는 현황계층에 축적되어 있는 데이터들 중 필요한 데이터 추출, 기존 사례 중 활용할 수 있는 모델 선정, 경상남도 차원에서의 역량, 공공데이터 활용 등의 관점에서 고려가 필요함



(그림10) 빅데이터 사업 추진을 위한 분석 계층

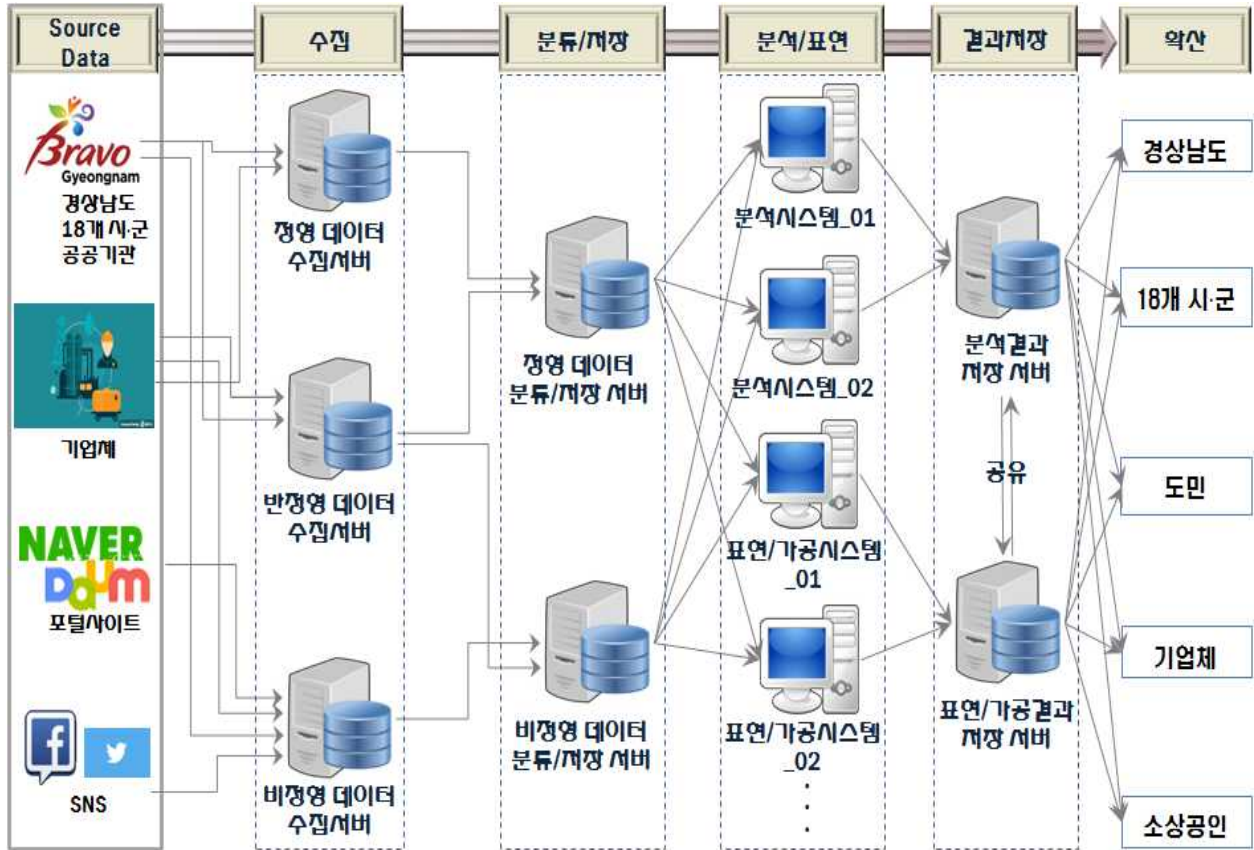
## 5) 경상남도 빅데이터 사업 추진 플랫폼

- 경상남도 빅데이터 사업의 효율적 추진을 위해서는 다양한 Activity를 지원할 수 있는 플랫폼이 마련되어야 할 것임
- 경상남도 빅데이터 사업 추진 플랫폼은 경상남도, 18개 시·군 및 관련 공공기관에 축적되어 있는 데이터, 문서, 민원데이터와 산업체 관련데이터, 소셜 & 뉴스데이터 등을 기반으로 함
- 이상의 기초데이터들을 수집하여 분류 및 저장하고, 목적에 따른 기초데이터를 추출하고, 적합한 분석기법을 도입하여 분석을 실시함
- 분석결과는 연구원 및 분석가와 같은 전문가들에 의해 해석하고, 결과는 분야별로 저장될 수 있도록 하여야 할 것임
- 이러한 결과들은 추후 효과적인 검색이 가능한 체계를 갖추어야 할 것이고, 나아가 시기별 Insight를 도출할 수 있는 기반을 고려해야 할 것임
- 따라서 경상남도 빅데이터 사업을 추진하기 위한 플랫폼 아키텍처를 정리하면 그림11과 같음



(그림11) 경상남도 빅데이터 사업 추진을 위한 플랫폼 아키텍처

- 빅데이터 사업 추진을 위한 플랫폼이 제대로 기능하기 위해서는 소스데이터 수집부터 활용까지 각 부문별 역할을 담당할 수 있는 하드웨어 및 소프트웨어가 필요한데 도식화하면 그림12와 같음



(그림12) 경상남도 빅데이터 사업 추진을 위한 플랫폼 아키텍처

- 먼저 공공기관, 기업체, 웹 등에서 수집할 수 있는 기초데이터는 다양하고, 매우 방대하므로 데이터 수집을 위한 시스템, 수집된 데이터의 저장 및 지속적으로 축적될 수 있는 규모의 하드웨어가 요구됨
- 수집된 기초데이터는 빅데이터 분석을 고려한 정형, 비정형 데이터로 구분하여 저장할 수 있는 하드웨어가 필요하고, 여기에 저장되어 있는 데이터를 활용하여 목적에 따른 분석 및 표현 작업을 위한 시스템이 구축되어야 할 것임
- 분석 및 표현된 결과는 추후 활용할 수 있도록 다른 서버에 저장될 수 있는 체계를 갖추고, 결과 저장 서버는 경상남도 및 18개 시·군, 도민, 기업체 및 소상공인 등에 서비스할 수 있는 기능을 포함함
- 아울러 각 단계(수집, 분류, 분석, 저장, 서비스 등)에서는 하드웨어를 컨트롤



룰할 수 있는 소프트웨어가 갖추어져야 할 것임

- 각 단계별 시스템 구축을 위한 하드웨어와 소프트웨어를 정리하면 표12와 같음

(표12) 빅데이터 사업 추진을 위한 단계별 H/W & S/W

| 사업수행단계  | H/W & S/W   | 비고         |
|---------|---|------------|
| 수집      | <ul style="list-style-type: none"> <li>○ HTTP/ODBC Connector</li> <li>○ Web crawler</li> </ul>  |            |
| 저장/관리   | <ul style="list-style-type: none"> <li>○ 대용량 데이터 저장 및 관리시스템 &amp; H/W</li> </ul>  | Tera-byte급 |
| 분석      | <ul style="list-style-type: none"> <li>○ 비정형/반정형 데이터 분석 Tool &amp; 저장 하드웨어(서버)</li> <li>○ 정형 데이터 분석 Tool &amp; 저장 하드웨어(서버)</li> <li>○ 데이터마이닝 S/W &amp; H/W</li> </ul> |            |
| 표현      | <ul style="list-style-type: none"> <li>○ 분석결과 표현(Visualize) Tool</li> </ul>   |            |
| 서비스(확산) | <ul style="list-style-type: none"> <li>○ 서비스 형태에 따른 S/W</li> </ul>  |            |

- 경상남도가 효과적인 빅데이터 사업을 추진하기 위해서는 각 단계별 다양한 S/W 및 H/W가 필요한데 사업의 규모, 추진 방식, 예산 등을 고려한 플랫폼이 구축되어야 할 것임

## 6) 경상남도 빅데이터 사업 추진 방향 종합

- 경상남도 빅데이터 사업 추진 방향은 정부 3.0, 경상남도 도정방침 및 핵심공약, 경상남도 현안 및 이슈, 실무전문가 대상 설문조사 등을 기반으로 초기단계에 적합한 사업 추진 방향(빅데이터 활용체계 구축, 전문성 확보, 플랫폼 구축 및 맞춤형 서비스 제공)을 설정함
- 핵심 방향에 따른 세부추진 방향도 경상남도의 현재 수준에서 시급히 갖추어 나가야 할 방향을 설정함
- 2016년부터 시범사업을 추진할 계획에 있는데, 시범사업 이후에는 향후 지속적인 빅데이터 서비스 제공을 위한 기반(예: 플랫폼 구축, 전담조직, 하드웨어 및 소프트웨어, 기초데이터 등)이 마련되어야 하고 이에 대한 보다 구체적인 연구가 필요함

## IV

## 경상남도 빅데이터 시범사업 추진 방안

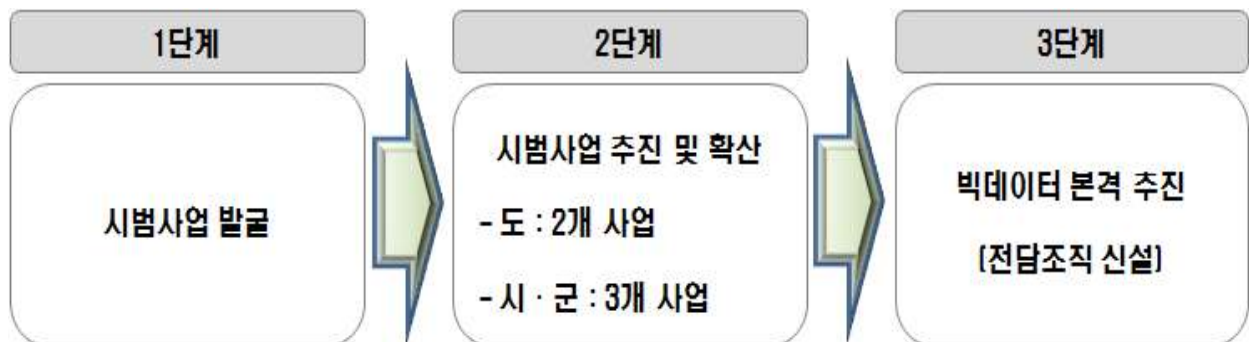
## 1. 경상남도 빅데이터 시범사업 추진 단계

- 빅데이터 분석은 공공기관, 기업체 및 시민들의 필요에 따라 다양한 형태로 분석할 수 있으며, 분석 결과를 기반으로 각 주체별 맞춤형 정책 및 서비스를 발굴할 수 있어 불확실한 미래에 대해 효과적으로 대비할 수 있는 장점이 있음
- 이러한 장점으로 인해 국내·외 공공기관 및 기업체에서는 삶의 질 개선, 경쟁력 강화 등을 목표로 빅데이터를 활용하고 있고, 확대해 나가고 있는 상황임
- 빅데이터는 공공분야 뿐만 아니라 기업의 더 나은 서비스 제공을 위한 기준 제시 및 새로운 가치 창출을 위한 기회로 인식되고 있음
- 따라서 아직 빅데이터의 활용이 미흡한 경상남도에서는 빅데이터의 필요성을 인식하고, 이와 관련된 도 정책 및 세부 사업을 조속히 추진할 필요가 있고, 단계별로 성과 확산, 시스템 및 전문인력 확보 등의 과정이 수반되어야 할 것임
- 1단계로 도내 다양한 측면에서 빅데이터 분석이 필요한 분야를 선정하고, 각 분야별 필요한 사업 아이템을 발굴하는 것이 선행되어야 할 것임
- 발굴된 사업 아이템 중 중요도 및 우선순위를 선정하고, 이를 기반으로 단계별 빅데이터 분석 사업을 추진하여 도민들의 삶의 질 개선 및 경상남도 경쟁력을 강화해 나가는 것이 바람직할 것으로 판단됨
  - 경상남도 차원에서 빅데이터 분석이 필요한 분야 선정
  - 다양한 분야 중 빅데이터 분석을 우선적으로 시행할 필요가 있는 세부 사업 선정

- 우선 추진 사업을 고려한 예산확보 및 최소한의 조직 체계 구축
  - 사업 추진 및 지속적인 모니터링
  - 사업 성과 극대화
- 2단계에서는 빅데이터 관련 사업의 성과를 입증하고, 획득된 데이터의 지속성 확보, 신규 사업 아이템 발굴, 시스템 및 전문인력 확대 등의 과정이 필요함
- 3단계로는 기 축적된 데이터를 활용하여 지역 현안에 신속히 대응할 수 있는 체계를 구축하고, 다양한 사업에 빅데이터 분석결과가 기준을 제시해 줄 수 있는 체계를 구축하는 것이 요구됨
- 이러한 과정을 통해 경상남도 도민들의 삶의 질 제고, 예산의 효율성 제고, 지역 경쟁력 확보 등이 가능할 것임

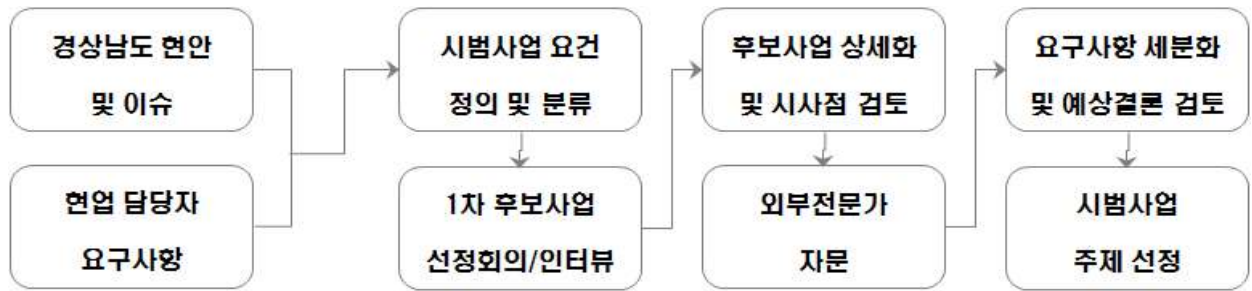
## 2. 경상남도 빅데이터 시범사업 추진 방향

- 경남미래 50년 전략사업, 도민복지 지원 및 사회안전망 구축 등을 고려하여 빅데이터 도입의 필요성을 인식하고 있고, 이를 통해 도민 맞춤형 공공서비스 정책개발에 활용할 계획에 있음
- 이를 위해 크게 3단계로 나누어 1단계에는 시범사업 추진, 2단계에는 사업 확산, 3단계는 본격 추진을 검토하고 있음



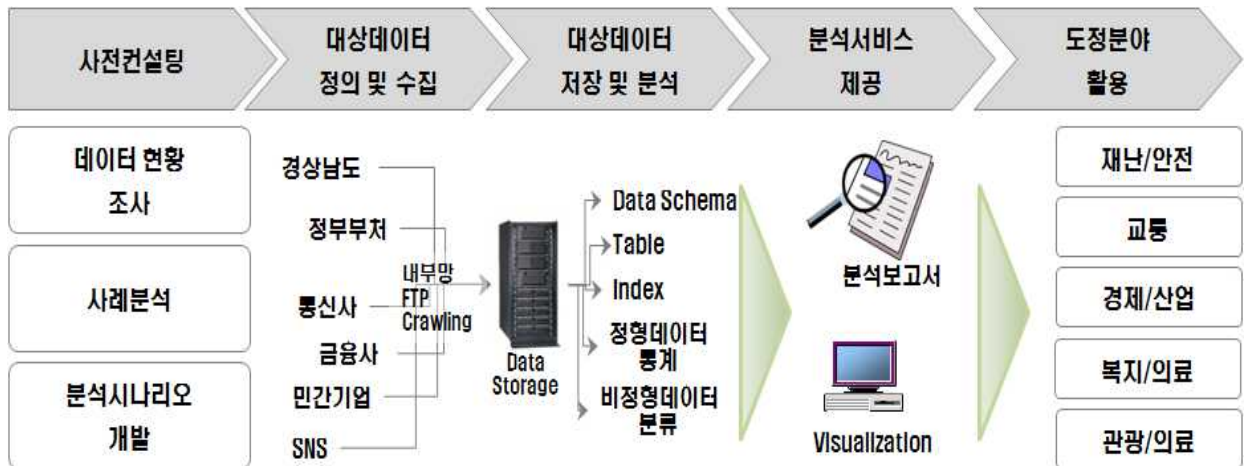
(그림13) 경상남도 빅데이터 사업 추진 단계별 전략(안)

- 특히 경상남도는 1단계의 효과적인 추진을 위해 현안 및 이슈, 현업 담당자의 요구사항을 파악하고, 시범사업 요건, 1차 후보사업 선정에 위한 인터뷰, 후보 사업군의 상세화, 자문회의 등의 과정을 거쳐 시범사업 분야 및 적용 가능 사업을 선정함



(그림14) 경상남도 빅데이터 1단계 시범사업 주제선정 절차

- 또한 2016년으로 예상하고 있는 2단계에서는 사업의 확산을 통해 경상남도 정책에 적극 활용할 계획을 가지고 있는데 이를 위해 사전컨설팅, 대상 데이터 수집, 저장 및 분석, 서비스 제공 및 정책 발굴 등의 과정을 고려하고 있음



(그림15) 경상남도 빅데이터 2단계 사업 추진 절차

- 이러한 과정을 거쳐 경상남도 정책 추진의 과학적 근거를 마련하고, 장기적으로 다양한 분야의 정책 발굴, 시행, 검증 등의 과정에서 기준을 제시할 수 있는 체계를 구상하고 있음

### 3. 경상남도 빅데이터 시범사업 적용 분야

○ 경상남도에서 선정한 시범사업 추진 분야는 10개 분야(도 : 5개 분야, 시·군 : 5개 분야)임(표13 참조)

- 도(5개 분야) : 도로, 소방, 주거, 감사, 농촌

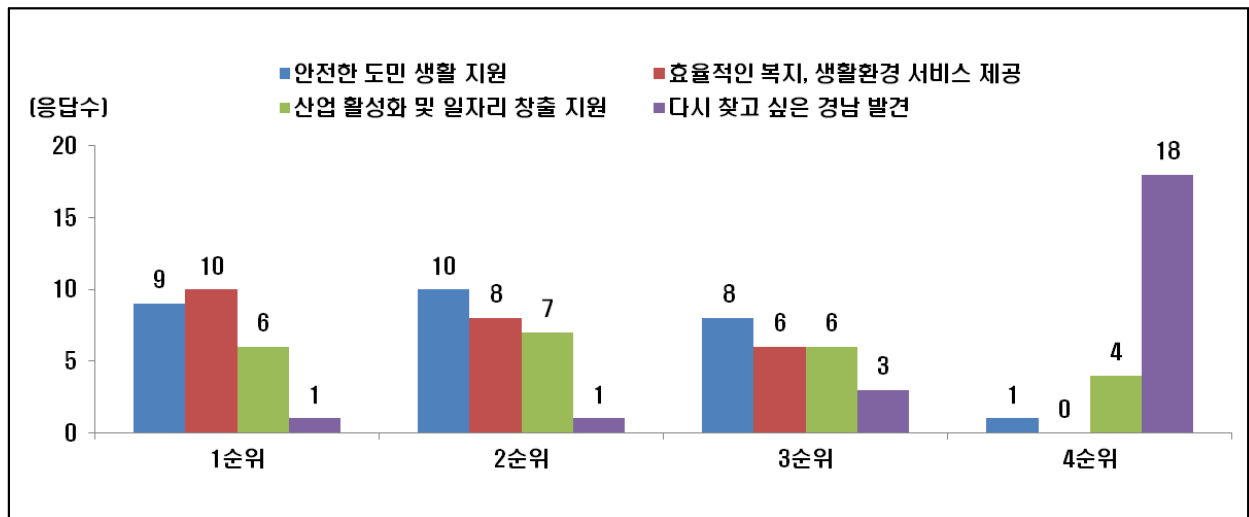
- 시·군(5개 분야) : 재난, 교통, 문화, 관광, 소상공인

(표13) 경상남도 빅데이터 적용가능 분야

| 구분  | 서비스 영역          | 사업명/주요내용   | 관련부서       |
|-----|-----------------|--|------------|
| 도   | 감사행정            | ○ 빅데이터를 활용한 새로운 감사대상·기법 발굴<br>- 지적사례, 신문, 민원 등 상관분석으로 새로운 패턴, 감사대상 발굴<br>- 데이터 기반의 감사로 감사 사각지대 해소                          | 감사관실       |
|     | 재난/안전<br>·교통    | ○ 도로환경 모니터링 및 분석으로 신속한 도로 보수<br>- 네비게이션의 도로 상태 체크 기능을 탑재하여 도로 데이터 수집·분석  | 도로과        |
|     | 주거행정            | ○ 공동주택 관리비·계약형태 분석을 통한 부조리 방지<br>- 공동주택 관리 등 집행내역 분석으로 적정여부 확인<br>- 공동주택 공사 및 용역업체 입찰정보 분석으로 공정성 확보                        | 건축과        |
|     | 농촌행정            | ○ 농산물 성장환경 빅데이터 분석을 통한 초보 농업인 영농활동 지원<br>- 농작물 재배 시 최적의 성장환경 및 생산량 정보 분석<br>- 최적의 성장환경 분석으로 작물 매출량 확대 및 경영비 절감             | 농업<br>정책과  |
|     | 재난/안전           | ○ 구급차 운영현황 분석으로 골든타임 확보 방안 마련<br>- 응급구조 활동 빈도와 위치정보, 응급실 도착시간과 이동거리 등을 분석하여 최적화된 응급구조 모델 구현<br>- 응급구조 수요가 많은 지역에 구급차 현장 배치 | 소방본부       |
| 시·군 | 교통              | ○ 출·퇴근 및 심야시간 버스이용 실태분석으로 버스노선 조정방안 마련<br>- 승하차 등 유동인구 분석<br>- 버스이용 수용에 맞게 신규 노선발굴 및 증감차 계획 수립                             | 교통부서       |
|     | 재난/안전           | ○ CCTV 설치 실태 분석으로 설치 및 관리 최적화 방안 마련<br>- 성별·연령별 유동인구 데이터, 범죄 HOT SPOT, 지리정보 등을 활용한 빅데이터 분석<br>- 최적의 CCTV 설치 모델 개발          | 정보화<br>부서  |
|     | 경제/산업<br>·문화/관광 | ○ 도내 축제 패턴 분석으로 지역경제 시너지 창출<br>- 다양한 빅데이터 기법을 활용한 축제 영향도 분석<br>- 다양한 지역 특성에 맞는 축제 활성화 전략 수립                                | 관광부서       |
|     | 문화/관광           | ○ 경남도내 외국이 관광 통계분석으로 국가별 맞춤형 관광정책 지원<br>- 경상남도 방문 외국인 대상 이동 동선, 소비패턴 분석<br>- 분석결과를 활용하여 외국인 국적별 맞춤형 관광정책 마련                |            |
|     | 경제/산업           | ○ 소상공인을 위한 빅데이터 분석 서비스 제공<br>- 상권, 인구 등의 분석을 통해 소상공인 창업 시 맞춤형 컨설팅 서비스 제공   | 경제산업<br>부서 |

## 4. 빅데이터 시범 사업 추진을 위한 설문조사

- 빅데이터 시범 사업의 효율적인 추진을 위해 경상남도 빅데이터 추진 방향 조사 시 함께 조사함
- 시범사업 추진 분야는 경상남도 도정방침, 현안 및 이슈, 기 경상남도에서 선정한 시범사업 추진분야 등을 고려하여 안전, 복지, 산업, 관광 분야로 구분하여 우선순위에 대해 조사하고, 각 분야별 세부사업 중 중점을 두어야 할 부분에 대해 조사를 실시함
- 시범사업 추진을 위해 안전, 복지, 산업, 관광 분야에 대해 우선순위를 조사한 결과 복지와 안전 분야에 초점을 맞추어야 한다는 의견이 가장 많았음

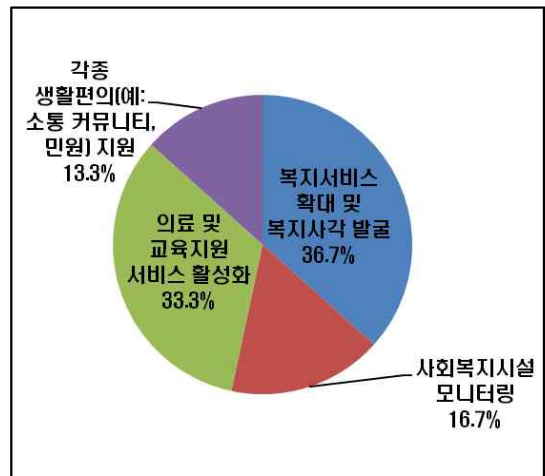


(그림16) 빅데이터 사업 추진 분야(다중 응답)

- 복지분야와 관련해서는 복지서비스의 확대 및 복지사각 지대를 발굴하는 것에 초점을 맞추어야 한다는 의견이 가장 많은 비중(36.7%)을 차지하는 것으로 분석되었고, 다음으로 의료 및 교육지원 서비스 활성화(33.3%), 사회복지 시설 모니터링(16.7%), 각종 생활편의(예: 소통 커뮤니티, 민원) 지원(13.3%)의 순으로 조사됨

(표14) 효율적인 복지, 생활환경 서비스 제공

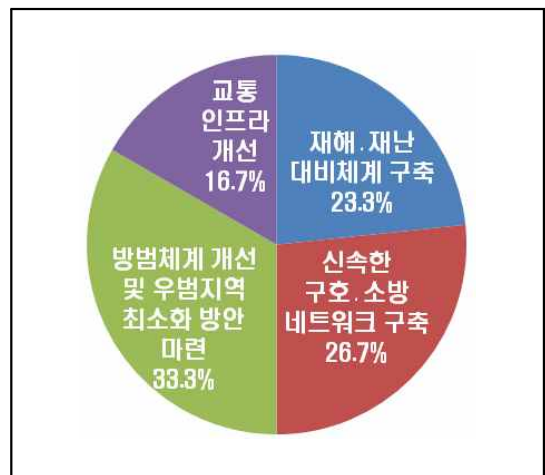
| 구분                         | 빈도 | 비율    |
|----------------------------|----|-------|
| 합계                         | 30 | 100.0 |
| 복지서비스 확대 및 복지사각 발굴         | 11 | 36.7  |
| 사회복지시설 모니터링                | 5  | 16.7  |
| 의료 및 교육지원 서비스 활성화          | 10 | 33.3  |
| 각종 생활편의(예: 소통 커뮤니티, 민원) 지원 | 4  | 13.3  |



○ 안전과 관련해서는 생활안전과 관련이 있는 방법체계 개선 및 우범지역 최소화 방안 마련에 초점을 맞추어야 한다는 의견(33.3%)이 가장 많았고, 다음으로 신속한 구호·소방 네트워크 구축, 재해·재난 대비체계 구축, 교통 인프라 개선의 순으로 조사됨

(표15) 안전한 도민 생활 지원

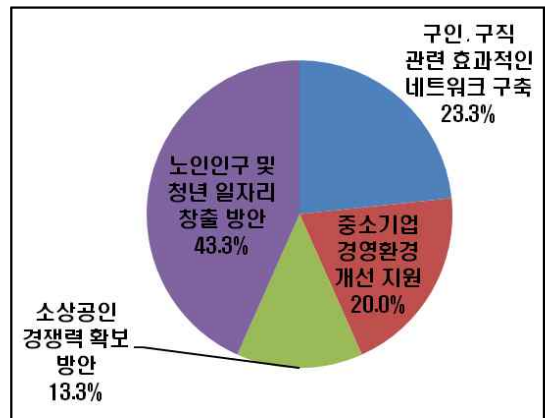
| 구분                       | 빈도 | 비율    |
|--------------------------|----|-------|
| 합계                       | 30 | 100.0 |
| 재해·재난 대비체계 구축            | 7  | 23.3  |
| 신속한 구호·소방 네트워크 구축        | 8  | 26.7  |
| 방법체계 개선 및 우범지역 최소화 방안 마련 | 10 | 33.3  |
| 교통 인프라 개선                | 5  | 16.7  |



○ 산업 활성화 및 일자리 창출 지원 부문에 있어서는 노인인구 및 청년 일자리 창출에 중점을 두어야 한다는 의견이 43.3%로 가장 많았고, 다음으로 구인·구직 관련 효과적인 네트워크 구축, 중소기업 경영환경 개선 지원, 소상공인 경쟁력 확보 방안의 순으로 조사됨

(표16) 산업 활성화 및 일자리 창출 지원

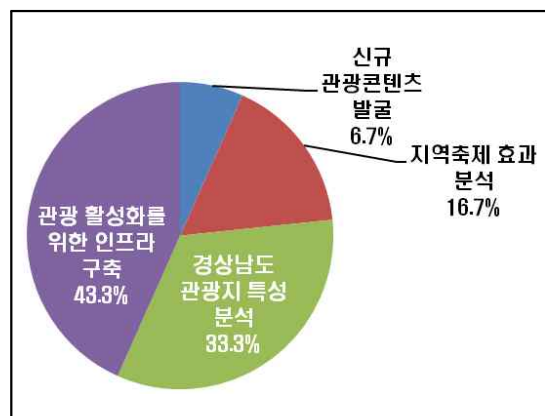
| 구분                    | 빈도 | 비율    |
|-----------------------|----|-------|
| 합계                    | 30 | 100.0 |
| 구인·구직 관련 효과적인 네트워크 구축 | 7  | 23.3  |
| 중소기업 경영환경 개선 지원       | 6  | 20.0  |
| 소상공인 경쟁력 확보 방안        | 4  | 13.3  |
| 노인인구 및 청년 일자리 창출 방안   | 13 | 43.3  |



- 끝으로 관광분야와 관련해서는 관광활성화를 위한 인프라 구축에 초점을 맞추어야 한다는 의견(43.3%)이 가장 많았고, 다음으로 경상남도 관광지 특성 분석, 지역축제 효과 분석, 신규 관광콘텐츠 발굴의 순으로 조사됨

(표17) 다시 찾고 싶은 경남 발견

| 구분                | 빈도 | 비율    |
|-------------------|----|-------|
| 합계                | 30 | 100.0 |
| 신규 관광콘텐츠 발굴       | 2  | 6.7   |
| 지역축제 효과 분석        | 5  | 16.7  |
| 경상남도 관광지 특성 분석    | 10 | 33.3  |
| 관광 활성화를 위한 인프라 구축 | 13 | 43.3  |



- 이상의 조사내용을 기반으로 경상남도 시범사업 추진 아이টে을 결정하고, 향후 시범 사업 추진을 통해 성과를 거둘 수 있도록 함

## 5. 경상남도 빅데이터 시범사업(안)

- 경상남도가 2016년 추진하게 될 시범사업으로는 기 조사·분석된 타 지자체 사례, 경상남도 도정방침, 현안 및 이슈, 사업추진방향 및 사업추진 분야 조사결과, 실무부서 발굴 아이টে, 빅데이터 시범사업 선정을 위한 경상남도



정보화 위원회 등의 내용을 기반으로 시범사업(안)을 마련함

- 시범사업(안)은 복지(2), 안전(1), 관광(2), 기타(1)로 총 6개 사업을 선정했고, 2016년부터 추진할 계획에 있음

## 1) 빅데이터를 이용한 사회복지시설 운영 개선 방안 발굴

### □ 목적

- 본 사업의 목적은 사회복지 대상(노인, 장애인, 여성, 다문화 등)별·지역별 사회복지시설의 적정성(규모, 종류, 위치 등)의 분석, 사회복지시설 관련 복지 누수 및 부정수급 유형 분석에 초점을 맞추고 있음
- 이외에도 공공요금 체납데이터, 개인신용정보 등을 기반으로 복지사각 대상자도 발굴할 계획임
- 이를 통해 적정 복지시설 제공 및 개선 방안을 마련하고, 복지누수 재발방지, 복지사각 대상자를 위한 복지서비스 제공 방안을 마련하는데 본 사업의 목적이 있음

### □ 주요내용

- 경상남도 및 관계기관이 보유하고 있는 데이터를 기반으로 도내 4,254개 사회복지시설의 과부족을 검토함
  - 과잉인 경우 통합 및 조정 방안 검토
  - 부족인 경우 신규 시설 조성 방안 검토
- 사회복지시설의 지도·단속 데이터를 기반으로 시설별·지역별 복지누수 및 부정수급 유형을 분석함
- 사회복지시설 지도·단속 시 시설유형별·지역별 등으로 구분하여 주요 체크포인트를 마련함
- 경상남도 복지서비스 현황 및 개인관련 정보 분석을 통한 복지사각대상을

발굴함

- 복지사각 대상 발굴 후 복지사각 대상별 기존 서비스 범위에 포함 또는 신규 서비스 제공 방안을 검토함

□ 단계별 추진사항

- 빅데이터를 이용한 사회복지시설 운영 개선 방안 발굴 사업의 추진은 데이터 수집, 분석방향 설정 및 분석, 정책발굴의 단계로 구분하여 추진하고, 각 단계별 추진사항을 정리하면 표18과 같음

(표18) 빅데이터를 이용한 사회복지시설 운영 개선 방안 발굴 사업 단계별 추진사항

| 사업수행단계       | 수집데이터 및 내용  | 비고             |
|--------------|---|----------------|
| 데이터 수집       | ○ 복지시설 현황 데이터<br>○ 복지시설 감사이력 데이터<br>○ 복지서비스 현황 데이터<br>○ 도시가스, 수도, 전기, 건강보험료 체납데이터<br>○ 개인신상관련 데이터(연령, 거주지 등)<br>○ 복지관련 민원데이터                        | 경상남도 및 18개 시·군 |
|              | ○ 개인신용정보  | 신용평가기관         |
|              | ○ SNS 데이터   | Web            |
|              | ○ 이동통신 데이터(유동인구 등)  | 이동통신사          |
| 분석방향 설정 및 분석 | ○ 분야별·지역별 사회복지시설 적정성(예: 규모, 종류, 위치 등) 분석<br>○ 사회복지시설 관련 복지 누수 및 부정수급 유형 분석<br>○ 복지사각 대상자 파악<br>○ 시간대별 복지사각 대상자 집중 지역 분석<br>○ 복지시설 및 복지사각 관련 도민의견 분석 |                |
| 정책발굴         | ○ 사회복지시설 신규 조성 및 기존 시설 조정 방안<br>○ 사회복지시설 관련 복지에산의 효과적 집행 방안<br>○ 경상남도 복지시설 점검을 위한 체크포인트<br>○ 복지사각 대상자 지원 방안   |                |

- 데이터 수집은 경상남도 정보통계담당관실 및 경남발전연구원 빅데이터센터(가칭)가 중심이 되어 경상남도 및 18개 시·군, 신용평가기관, 이동통신사, SNS 등에서 데이터를 수집함

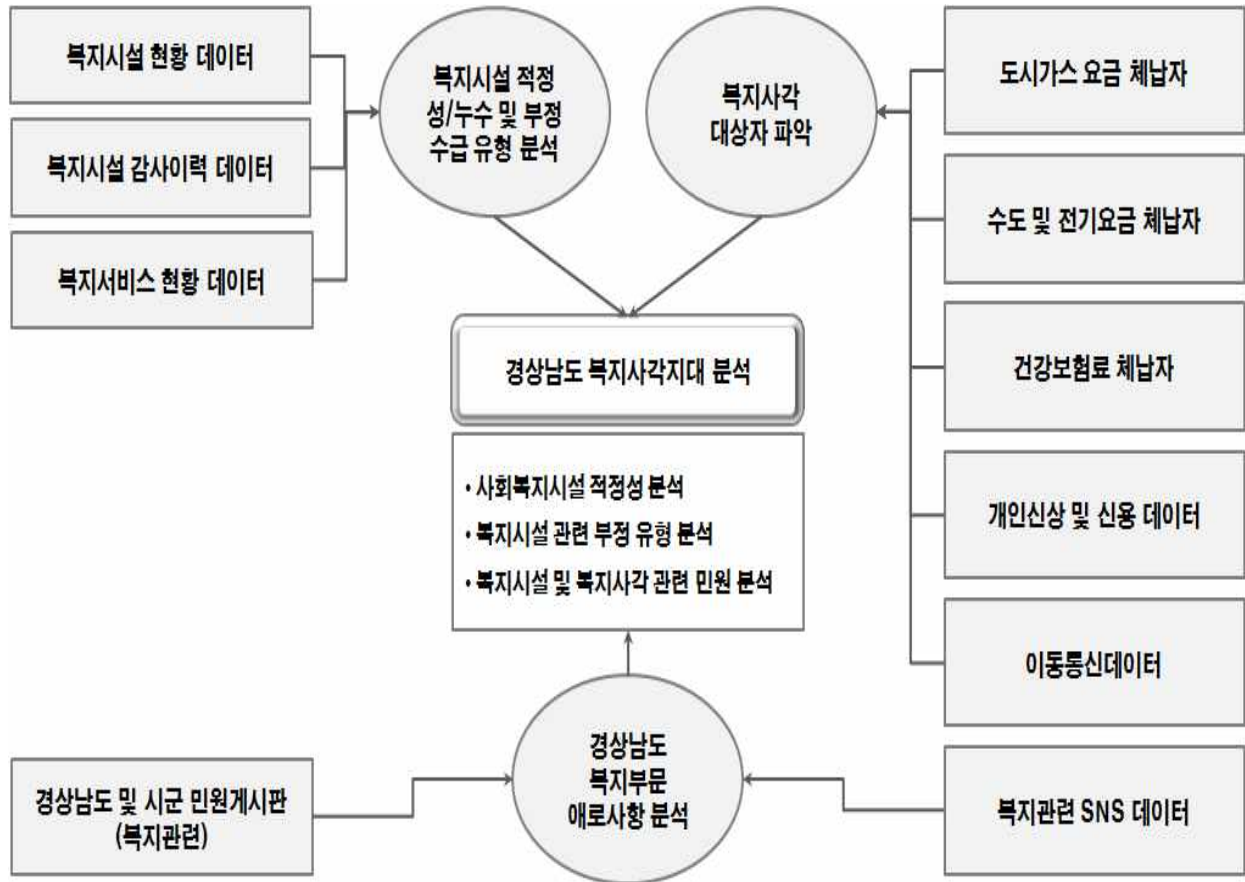
| Source Data   |  |   |   |
|---|--|---|---|
| 경상남도/18개 시·군  | 신용평가기관   | 이동통신사   | SNS   |
| <ul style="list-style-type: none"> <li>· 복지시설 현황 데이터</li> <li>· 복지시설 감사이력 데이터</li> <li>· 복지서비스 현황 데이터</li> <li>· 도시가스, 수도, 전기, 건강 보험료 체납데이터</li> <li>· 개인별 연령, 거주지 등</li> <li>· 복지관련 민원데이터</li> </ul> | <ul style="list-style-type: none"> <li>· 개인신용정보</li> </ul> | <ul style="list-style-type: none"> <li>· 평균이동거리/위치</li> <li>· 통신요금 사용량</li> </ul> | <ul style="list-style-type: none"> <li>· 트위터</li> <li>· 페이스북 등</li> </ul> |

경상남도(자료요청)/GNDI(기관간 연계, 취합, 관리 등)

|           |  |
|-----------|--|
| 데이터 수집 방안 | <ul style="list-style-type: none"> <li>· 복지대상 개인정보 수집시 다양한 측면에서의 검토 필요(예: 개인정보 보호법, 국민기초생활 보장법, 기초노령연금법, 사회복지사업법, 포괄적 정보 수집 사전동의서의 징구 등)</li> <li>· 지역별 다양한 기초데이터 확보</li> <li>· On/Off Line 복합 수집 실시</li> <li>· 기존 복지서비스 및 시설 현황 자료 수집</li> <li>· 경상남도 SNS 데이터 확보 시 실적, 체납, 단수, 단전 등 실생활 애로를 고려한 키워드 중심의 데이터 Crawling</li> </ul> |
|-----------|--|

(그림17) 데이터 수집 방안

- 데이터 수집 시 개인정보 수집, 빅데이터 분석을 고려한 기초데이터 확보, 핵심 소셜데이터 수집 등이 고려되어야 할 것임
- 수집된 데이터를 기반으로 사회복지시설의 적정성, 복지누수 및 부정수급 유형 분석, 복지사각 대상자 파악, 복지시설 및 복지사각 관련 민원데이터 분석 등을 통해 경상남도 사회복지시설의 신규 조성 및 조정 방안, 복지에 산의 효과적 집행, 복지사각 대상자 지원 방안 등의 정책을 발굴함
- 이러한 분석 프로세스를 정리하면 그림18과 같음



(그림18) 분석 프로세스

## □ 기대효과

- 본 사업의 수행을 통해 경상남도 사회복지시설의 부정·부패 재발을 방지하고, 신규 복지제공 대상자 선정 및 복지서비스를 발굴할 수 있을 것으로 예상됨
- 아울러 경상남도 예산 중 많은 비중을 차지하고 있는 복지예산의 효과성을 제고 할 수 있을 것으로 기대됨

## 2) 노인계층 삶의 질 향상을 위한 빅데이터 분석

### □ 목적

- 본 사업을 통해서는 경상남도에 위치해 있는 노인계층의 분포와 이들을 위해 제공되는 복지서비스의 현황을 파악하고, 현황을 기반으로 경상남도 노인계층의 특성을 분석함

- 특성 분석을 통해 노인계층 생활전반을 지원할 수 있는 복지서비스(일자리, 안전시설, 의료시설 연계 등)를 발굴하는데 목적이 있음

## □ 주요내용

- 경상남도 시·군별 노인계층 분포 및 복지서비스 제공 현황을 파악하고, 이를 기반으로 빅데이터 분석 대상 지역(도시, 도시와 농촌, 농촌 등)을 선정함
- 노인계층 연령대별 취업 형태 및 희망일자리 분석을 통해 장래 노인일자리 창출을 위한 기초자료를 생성함
- 노인계층의 경우 뜻밖의 사고율이 높는데 사고유형, 지역 및 처리현황 등을 분석함
- 또한 지역별 노인계층의 질환 유형 분류 및 상황별·질환별 의료시설과의 연계 방안을 모색함
- 이외에도 빅데이터 분석을 통해 노인계층 지원을 위한 복지시설 및 서비스를 위한 아이템을 발굴하고자 함

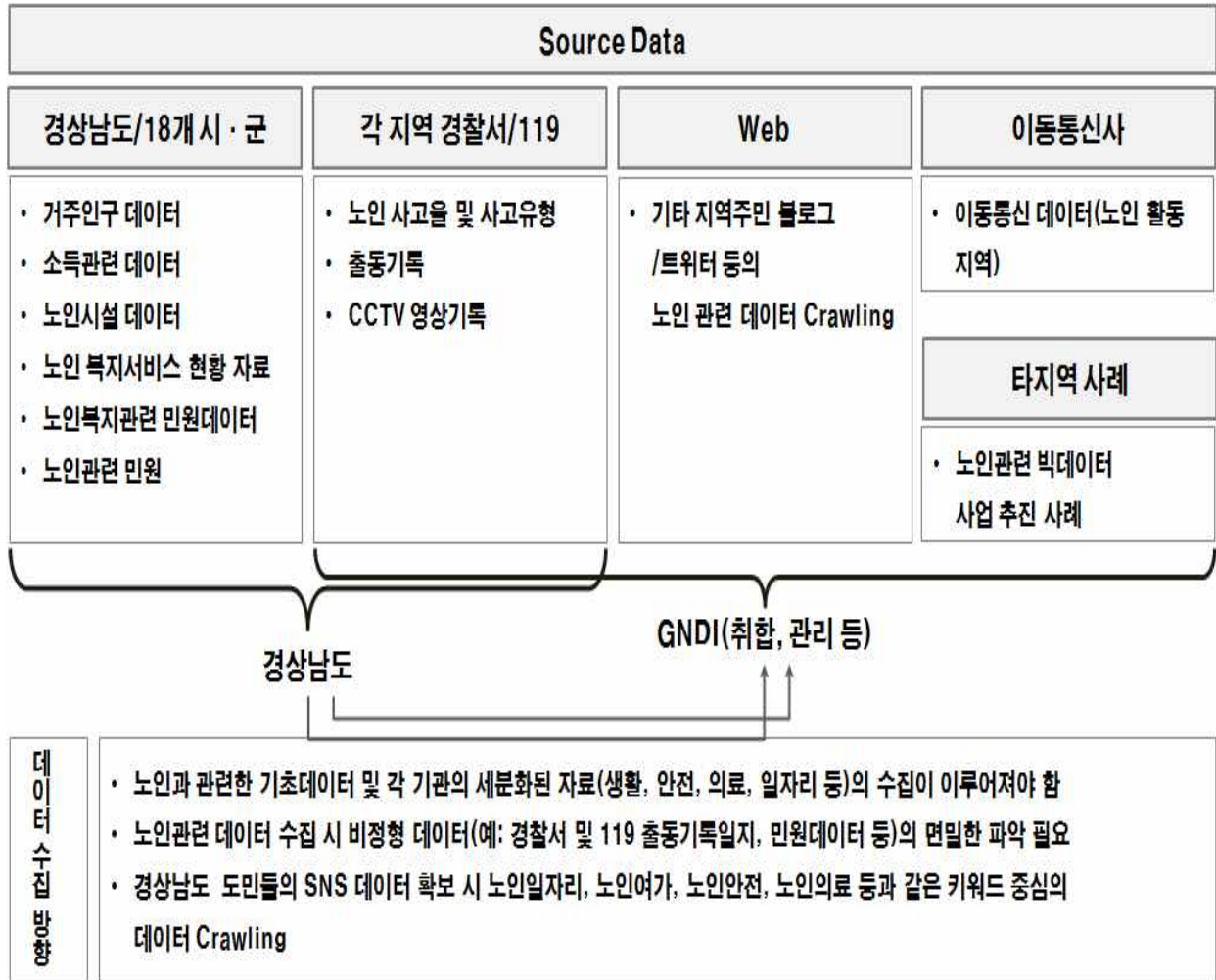
## □ 단계별 추진사항

- 노인계층 삶의 질 향상을 위한 빅데이터 분석 사업의 추진 시 각 단계별 추진사항을 정리하면 표19와 같음

(표19) 노인계층 삶의 질 향상을 위한 빅데이터 분석 사업의 단계별 추진사항

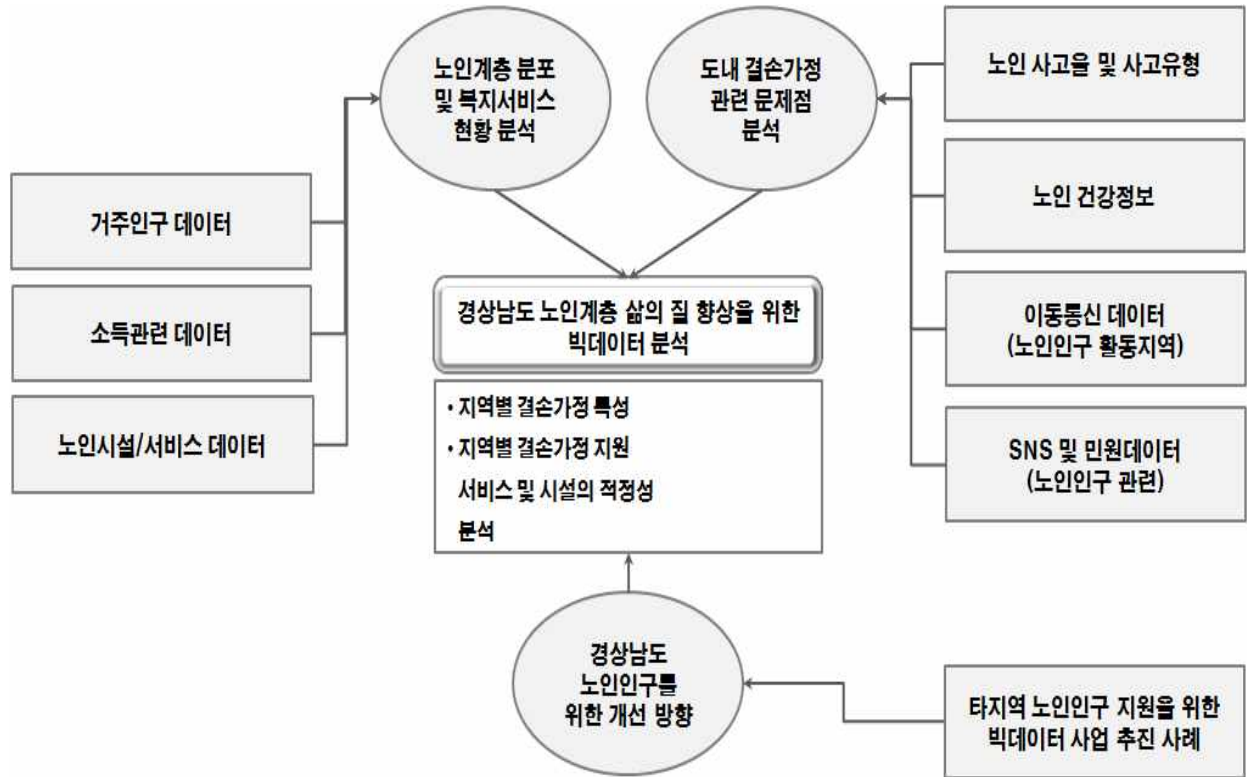
| 사업수행단계       | 수집데이터 및 내용  | 비 고             |
|--------------|---|-----------------|
| 데이터 수집       | <ul style="list-style-type: none"> <li>○ 거주인구 데이터</li> <li>○ 소득관련 데이터</li> <li>○ 노인시설 정보</li> <li>○ 노인 복지서비스 현황 자료</li> <li>○ 노인복지관련 민원데이터</li> </ul>   | 경상남도 및 18개 시·군  |
|              | ○ 노인 사고율 및 사고유형   | 각 지역 경찰서/119구조대 |
|              | ○ 노인 건강정보   | 지역 보건소          |
|              | ○ 이동통신 데이터(노인 활동지역 관련)  | 이동통신사           |
|              | ○ SNS 데이터   | Web             |
| 분석방향 설정 및 분석 | <ul style="list-style-type: none"> <li>○ 노인계층 분포 및 복지서비스 제공 현황 분석</li> <li>○ 지역별·시간대별 노인인구 분포현황 분석</li> <li>○ 노인인구 특성(취업형태, 사고율 및 사고유형, 질환 유형 등) 분석</li> </ul>  |                 |
| 정책발굴         | <ul style="list-style-type: none"> <li>○ 노인인구 특성 및 노인 희망일자리를 고려한 일자리 창출(예: 학교지킴이, 급식 도우미, 숲생태해설가 등) 정책 발굴</li> <li>○ 노인인구 지원을 위한 안전시설(예: 안전표지판, 과속방지턱, 횡단보도 등) 조성 계획</li> <li>○ 노인인구 관련 의료시설 확충 및 효과적인 연계 방안</li> <li>○ 신규 노인 복지 서비스 발굴 및 복지시설 조성 방안</li> </ul> |                 |

- 단순히 노인계층 복지시설 및 서비스에 국한된 것이 아니라 분석을 통해 종합적인 방안 마련이 핵심임
- 데이터 수집은 경상남도 정보통계담당관실 및 경남발전연구원 빅데이터센터(가칭)가 중심이 되어 경상남도 및 18개 시·군, 타지역 사례, SNS 등에서 데이터를 수집함



(그림19) 데이터 수집 방안

- 수집된 데이터를 기반으로 노인계층 분포 및 복지서비스 제공 현황 분석, 지역별·시간대별 노인인구 분포현황 분석, 노인인구 특성(취업형태, 사고율 및 사고유형, 질환 유형 등) 분석하여 향후 노인인구들을 지원할 수 있는 정책과 연계될 수 있도록 함
- 이러한 분석 프로세스를 정리하면 그림20과 같음



(그림20) 분석 프로세스

### □ 기대효과

- 본 사업을 통해 경상남도에 기 축적되어 있는 노인계층 관련 데이터의 체계를 효과적으로 재정립할 수 있음
- 노인계층을 위한 맞춤형 복지서비스 제공이 가능하고, 효과적인 복지예산의 활용을 통한 신뢰성 확보가 용이함

## 3) 빅데이터 분석을 통한 안전 사각지대 최소화 방안

### □ 목적

- 본 사업은 창녕군의 안전시설 설치 현황 및 사건·사고 현황을 파악하고, 이를 기반으로 문제점 및 안전사각지대를 파악함
- 안전과 관련한 문제점 및 안전사각지대를 최소화하기 위한 장비의 설치, 순찰 체계 정비 등과 같은 방안을 마련 것이 목적임



## □ 주요내용

- 창녕군 안전시설(예 : CCTV, 가로등, 표지판 등)의 설치 위치, 개수 등과 같은 현황을 파악함
- 경찰서 및 119구조대 등에 축적되어 있는 사건·사고 현황 데이터를 기초로 지역별·시기별·시간대별 안전사각지대를 파악함
- 사건·사고의 경우 유형을 분류하고, 유형별 사건·사고의 발생 패턴 분석을 통해 장래 불특정 사건·사고를 예측하는데 기반을 마련하도록 함
- 아울러 민원데이터 및 게시판 등과 같은 비정형 데이터 확보를 통해 안전관련 문제점 분석 및 신규 서비스를 발굴하고자 함

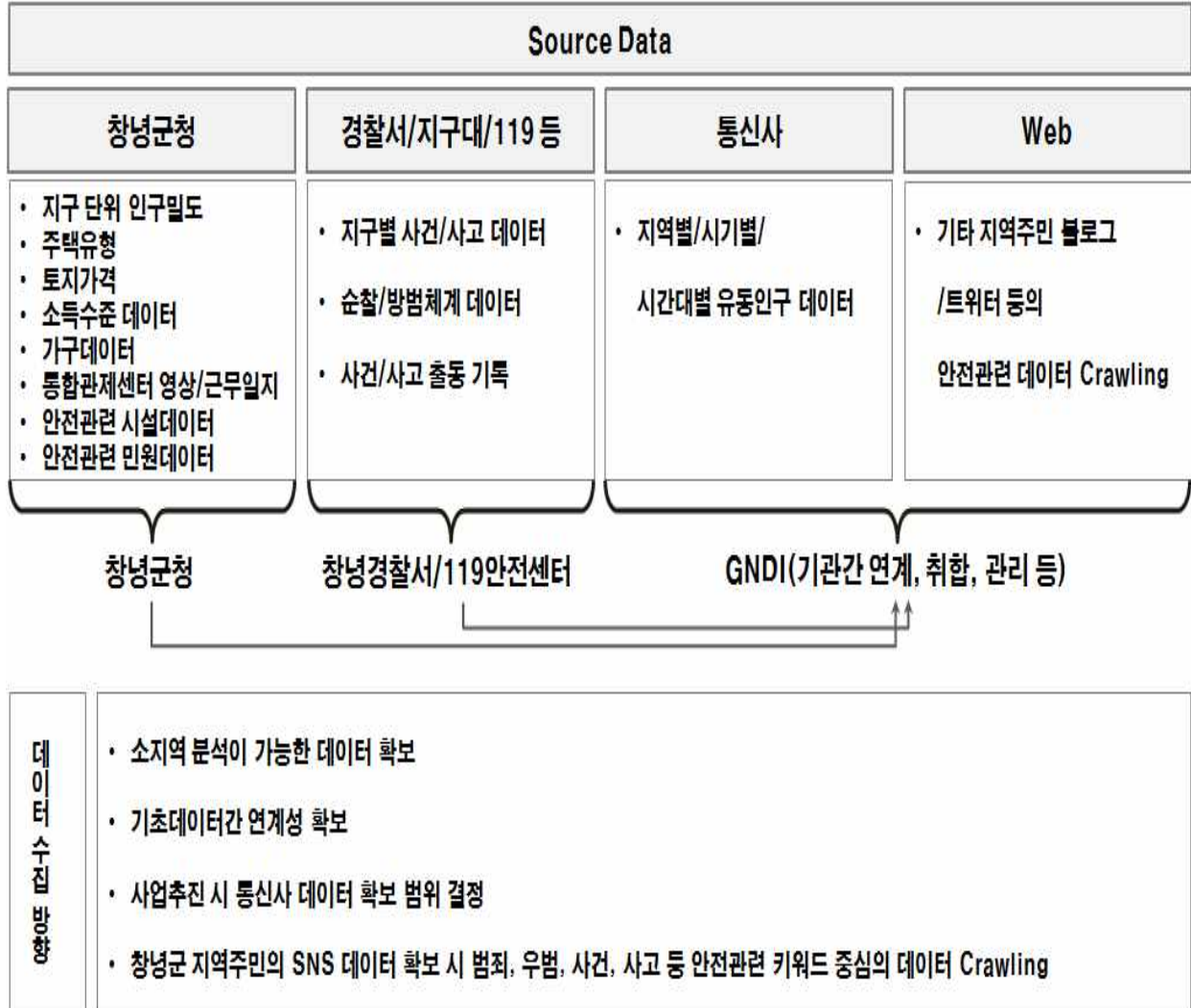
## □ 단계별 추진사항

- 빅데이터 분석을 통한 안전 사각지대 최소화 방안 마련을 위한 각 단계별 수행 내용을 정리하면 표20과 같음

(표20) 빅데이터 분석을 통한 안전 사각지대 최소화 방안 사업의 단계별 추진사항

| 사업수행단계       | 수집데이터 및 내용  | 비 고           |
|--------------|---|---------------|
| 데이터 수집       | <ul style="list-style-type: none"> <li>○ 소지역 단위 인구밀도</li> <li>○ 주택유형</li> <li>○ 토지가격</li> <li>○ 소득수준 데이터</li> <li>○ CCTV 설치 현황 데이터</li> <li>○ 안전관련 민원데이터</li> </ul> | 창녕군           |
|              | <ul style="list-style-type: none"> <li>○ 통합관제센터 근무일지</li> <li>○ 통합관제센터 영상데이터</li> </ul>   | 창녕군<br>통합관제센터 |
|              | <ul style="list-style-type: none"> <li>○ 사건·사고 데이터</li> </ul>   | 창녕경찰서/지구대 등   |
|              | <ul style="list-style-type: none"> <li>○ 119출동기록 데이터</li> </ul>   | 창녕소방서         |
|              | <ul style="list-style-type: none"> <li>○ 유동인구 데이터</li> </ul>  | 이동통신사         |
|              | <ul style="list-style-type: none"> <li>○ SNS 데이터(안전 관련)</li> </ul>  | Web           |
| 분석방향 설정 및 분석 | <ul style="list-style-type: none"> <li>○ 안전사각지대 선정 및 특성 분석</li> <li>○ 지역별 사건·사고 유형 분석</li> </ul>  |               |
| 정책발굴         | <ul style="list-style-type: none"> <li>○ 안전시설 추가 설치 계획 수립</li> <li>○ 안전사각지대 환경 개선 사업 추진</li> <li>○ 방법체계 개선 및 강화</li> </ul> <p>⇒ 안전 사각지대 개선을 위한 단계별 정책 추진</p>        |               |

- 데이터 수집은 창녕군청, 창녕경찰서/119안전센터 및 경남발전연구원이 중심이 되어 관련 자료를 수집함



(그림21) 데이터 수집 방안

- 수집된 데이터를 기반으로 안전사각지대 선정 및 특성 분석, 지역별 사건·사고 유형 분석을 실시하고, 향후 창녕군 안전사각지대 최소화를 위한 정책 발굴과 연계될 수 있도록 함
- 이러한 분석 프로세스는 그림22와 같음



(그림22) 분석 프로세스

#### □ 기대효과

- 안전관련 다양한 측면에서의 분석 활동을 통해 불의의 사건·사고를 미연에 방지할 수 있는 기반조성이 가능함
- 안전 사각지대 최소화를 위한 효과적인 모델 개발이 가능하고, 장래 경상남도 전역을 대상으로 확대·적용이 가능할 것으로 예상됨

### 4) 지역축제 시 카드사용 데이터 분석을 통한 관광정책 발굴

#### □ 목적

- 본 사업은 밀양시에서 개최되는 지역 축제와 관련된 빅데이터 분석을 실시하고자 함
  - 밀양시 지역별·시기별 개최 행사 현황 파악
  - 행사별 관람객 특징, 소비패턴 및 의견 분석
- 이를 통해 밀양시 행사 개최 시 효과적인 계획을 수립하는데 대한 기반을 마련하는데 목적이 있음

## □ 주요내용

- 밀양시에서 개최되는 행사 개최지 주변의 상권, 관광지 및 전통시장의 매출 데이터의 수집 및 분석
- 행사별 매출액, 관광객의 이동 경로 및 소셜데이터 등에 대한 빅데이터 분석을 통해 밀양시 행사의 취약점과 개선사항을 도출함
- 이외에도 밀양시에서 개최되는 행사별 고려사항을 정리하여 향후 행사 개최 시 기초자료로 활용할 수 있는 체계를 구축함

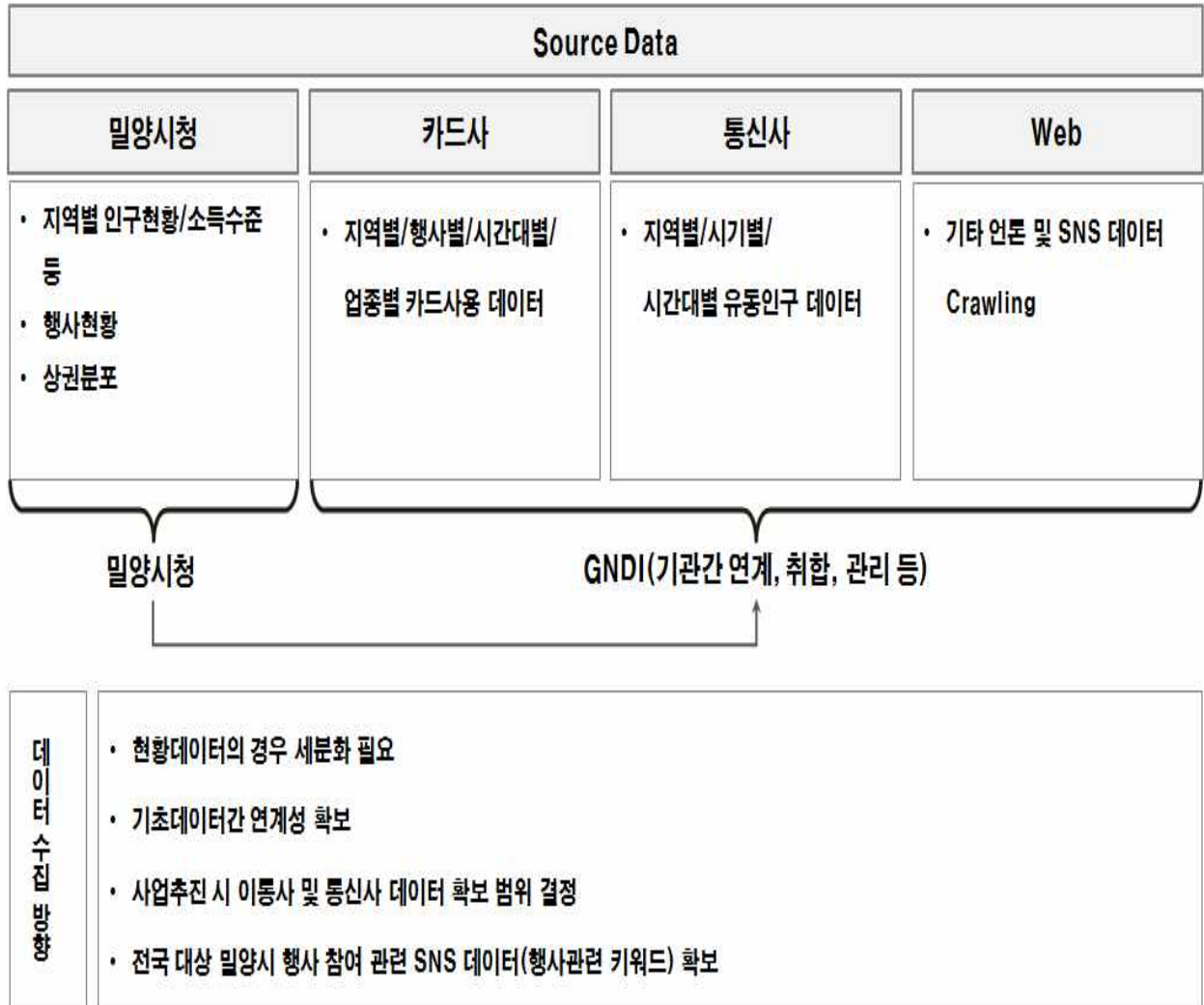
## □ 단계별 추진사항

- 밀양시 축제 개최 시 카드사용 데이터 분석을 통한 관광정책 발굴 사업의 단계별 추진사항은 표21과 같음

(표21) 지역축제 시 카드사용 데이터 분석을 통한 관광정책 발굴사업의 단계별 추진사항

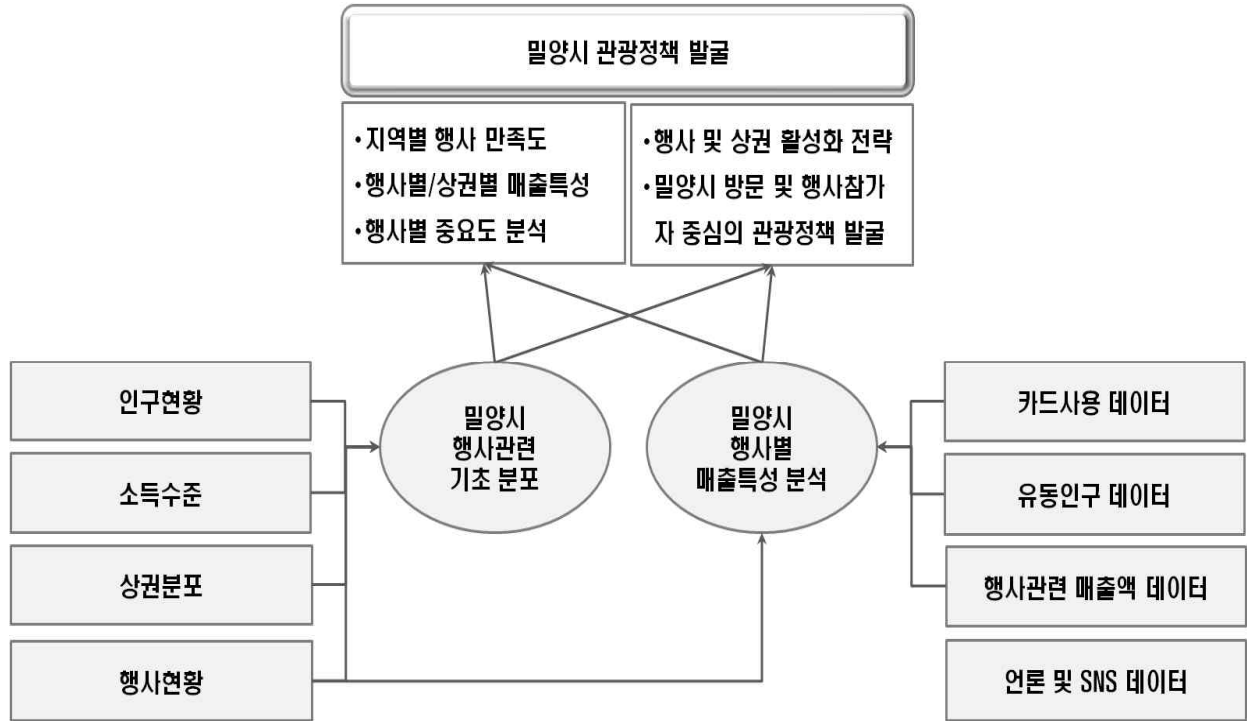
| 사업수행단계       | 수집데이터 및 내용                                      | 비 고   |
|--------------|---|-------|
| 데이터 수집       | ○인구현황 데이터<br>○지역별 상권현황 데이터<br>○시기별·지역별 행사현황 데이터 | 밀양시   |
|              | ○행사 기간별 카드사 데이터                                 | 카드사   |
|              | ○유동인구 데이터                                       | 이동통신사 |
|              | ○언론 및 SNS 데이터(밀양시 행사 관련)                        | Web   |
| 분석방향 설정 및 분석 | ○시기별·지역별 맞춤형 행사 전략(행사위치, 프로그램, 마케팅, 주요 대상 등)마련  |       |
| 정책발굴         | ○밀양시 행사별 세부전략 수립<br>○지역별·행사별 상권 활성화 전략 제시       |       |

- 데이터 수집은 밀양시와 경남발전연구원이 중심이 되어 관련 자료를 수집함



(그림23) 데이터 수집 방안

- 수집된 데이터를 기반으로 시기별·지역별 맞춤형 행사 전략(행사위치, 프로그램, 마케팅, 주요 대상 등)을 위한 기반을 마련함
- 이러한 분석 프로세스는 그림24와 같음



(그림24) 분석 프로세스

## □ 기대효과

- 카드사용 데이터 및 소셜데이터의 분석은 밀양시 행사별 맞춤형 전략을 마련할 수 있도록 하고, 지역 소상공인들의 매출 및 경쟁력을 강화하는데 도움을 줄 수 있음
- 장기적인 관점에서는 밀양시 행사별 타겟팅 전략 마련이 용이할 것으로 판단되고, 경상남도 타 지역 축제 효과 분석에도 활용할 수 있을 것임

## 5) 소셜데이터 분석을 통한 경상남도 관광정책 발굴

### □ 목적

- 경상남도의 종합적인 관광정책 발굴을 위해 우리나라 국민들의 블로그 및 SNS를 대상으로 경상남도 및 경상남도 관광지에 대한 인식을 분석함
- 아울러 경상남도를 방문하는 외국인들에 대해서도 페이스북, 트위터 등을 통해 경상남도 및 경상남도 관광지에 대한 인식이 어떠한지 빅데이터 분석을 실시함

- 이러한 분석을 토대로 경상남도의 관광명소를 발굴하고, 관광객 유치에 위한 전략 수립의 기반을 마련하는데 본 사업의 목적이 있음

## □ 주요내용

- 우리나라 국민들의 블로그, 페이스북, 트위터 등에서 데이터를 수집하여 경상남도를 방문했을 때의 긍정적인 측면과 부정적인 측면을 분석함
- 또한, 외국인 관광객들이 주로 참고하는 웹사이트 및 SNS 등의 데이터를 기초로 경상남도 방문 및 관광 시 장소와 이용시설별 긍정적인 측면과 부정적인 측면을 분석함
- 내·외국인의 분석 결과를 바탕으로 경상남도의 관광명소를 발굴하고, 관광지 시설 개선과 도로, 숙박시설, 쇼핑센터, 재래시장 등을 개선하기 위한 방안을 마련함
- 관광명소와 관련 인프라 등의 연관분석을 통해 지역별·시설별 개선 사항을 도출할 수 있도록 함

## □ 단계별 추진사항

- 내·외국 대상 소셜데이터 분석을 통한 경상남도 관광정책 발굴을 위한 단계별 추진사항은 표22와 같음

(표22) 소셜데이터 분석을 통한 관광정책 발굴 사업의 단계별 추진사항

| 사업수행 단계      | 수집데이터 및 내용  | 비 고        |
|--------------|---|------------|
| 데이터 수집       | ○내·외국인 블로그  | 각 부문별 Site |
|              | ○내·외국인 SNS  |            |
| 분석방향 설정 및 분석 | ○경상남도 관광관련 긍정 키워드 분석<br>○경상남도 관광관련 부정 키워드 분석  |            |
| 정책발굴         | ○관광 프로그램 개발<br>○문화시설 및 자연명소 개선 사업<br>○부정적인 키워드 개선을 위한 세부사업(예: 도로 확장, 숙박시설 확충, 재래시장 정비 등) 연계 |            |

## □ 기대효과

- 경상남도 각 지역별 맞춤형 관광정책 수립이 용이하고, 이로 인해 관련분야 효과적인 예산 집행이 가능함
- 관광정책을 비롯한 교통, 산업, 문화 등 다양한 분야의 연계 정책 발굴에도 도움이 될 수 있을 것으로 판단됨

## 6) 지역별 주부계층 소셜데이터 분석을 통한 생활밀착형 정책 발굴

### □ 목적

- 경상남도 각 지역별 엄마들이 활동하는 온라인 카페에서 공유되는 다양한 의견을 분석함
- 지역적 차원에서 추진되는 정책과 연계하여 도민이 체감할 수 있는 사업 및 정책을 발굴하는데 본 사업의 목적이 있음

### □ 주요내용

- 경상남도 지역별 엄마들이 의견 및 정보를 공유하고 있는 맘스카페 및 유사 카페를 대상으로 데이터를 수집함
- 지역별 맘스카페 공유 의견을 토대로 분야별로 분류하고, 사업 추진이 필요한 아이템을 도출함
- 도출된 사업아이템과 지역별 추진 정책과 상관분석을 실시하여 추진 가능성이 있는 아이템을 선별함
- 각 지역별 필요한 정책 및 사업들을 대상으로 우선순위를 선정하고, 우선 추진이 필요한 정책 및 사업에 대해서는 추진 방안까지 검토함

### □ 단계별 추진사항

- 지역별 주부계층 소셜데이터 분석을 통한 생활밀착형 정책을 발굴하기 위한 단계별 추진사항은 표23과 같음



(표23) 주부계층 소셜데이터 분석을 통한 생활밀착형 정책 발굴 사업의 단계별 추진사항

| 사업수행단계       | 수집데이터 및 내용  | 비 고       |
|--------------|---|-----------|
| 데이터 수집       | ○ 지역별 맘스 카페   | 각 카페 Site |
| 분석방향 설정 및 분석 | ○ 지역별 카페 내의 현안문제 발굴<br>○ 분야별 문제 분석<br>○ 지역별 추진 정책 및 카페 분석 내용 연계 |           |
| 정책발굴         | ○ 지역별 필요 정책 및 사업 도출<br>○ 도출된 정책 및 사업 대상 우선순위 선정                 |           |
| 활용방안         | ○ 경상남도 전 분야 정책 및 사업 발굴 기반                                       |           |

## □ 기대효과

- 각 지역별 실생활에서 애로를 겪고 있는 사항들이 도출될 수 있으므로 지역별 맞춤형 정책 수립이 용이할 것으로 판단됨
- 맘스카페 등은 생활 전반에 대한 내용들이 공유되므로 새로운 아이템 발굴 및 도민들이 체감할 수 있는 정책 및 사업발굴이 용이할 것으로 기대됨

## 6. 빅데이터 시범사업 추진을 위한 시스템 구축 방향

- 빅데이터 시범사업 추진을 위해서는 데이터 수집, 저장, 분석, 표현 등 여러 단계에 필요한 하드웨어와 소프트웨어가 요구됨
- 각 단계별 하드웨어와 소프트웨어의 사용도 다양하여 어떤 방향을 설정하는 것이 타당한지에 대한 사전 검토가 필요함
- 효과적인 시범사업 추진을 위한 시스템 구축을 위해 기존 빅데이터 시범사업 추진 경험이 있는 전문가들을 대상으로 인터뷰를 실시하고, 이를 기반으로 경상남도 빅데이터 시범사업 추진에 필요한 시스템 구축 방향을 설정함

### 1) 시스템 구축을 위한 전문가 인터뷰

#### □ 한국정보화진흥원

- 한국정보화진흥원에서는 K-ICT 빅데이터센터를 운영하고 있고, 신규 사업발굴, 예산확보, 사업발주 및 관리 등의 업무를 담당하고 있음

- 한국정보화진흥원에서 빅데이터센터를 운영한 경험에 비추어볼 때 지자체 차원에서의 빅데이터 분석을 위해서는 소규모로 시작해서 노하우를 쌓은 후 담당 업무를 고려한 솔루션 및 시스템을 구축하는 것이 효과적인 것이라고 제안함
  - OS 및 DB Server
  - 3~4 Client 시스템(PC Level)
- 경남발전연구원내 소규모 빅데이터 분석팀은 기초 데이터 관리, DB관리, 공공데이터를 활용한 분석, 분석결과 해석, 빅데이터 분석이 필요한 아이템 발굴 등이 초기에 담당할 수 있는 영역이므로 이를 고려한 시스템 구축이 적합할 것이라는 의견임
  - H/W : DB Storage 기능 서버 1대, Client PC 3~4대
  - S/W : DB Server(예: MS SQL, MySQL 등), OS 및 Open Application 등

## □ 해운대구청

- 부산시 해운대구청은 2013년 1월 빅데이터분석팀을 신설하고, 관광정책과 관련한 설문, 현황조사 등 실태파악을 실시함
  - 빅데이터 분석팀은 3명으로 구성
  - 빅데이터 분석이 필요한 계획 수립, 공공데이터 분석, 기초데이터 관리 및 DB화 등의 업무 담당
- 해운대구청은 2013년 5월~10월, 2014년 5월~10월까지의 트위터, 페이스북 및 블로그 등 SNS 데이터(1만 50건)를 활용한 빅데이터 분석을 실시하여 해운대 관광의 매력요인과 방해요인을 분석하고, 이에 따른 관광정책을 마련한바 있음
- 해운대구청 빅데이터분석팀의 의견은 연구원이라는 강점이 있으므로 빅데이터 분석이 필요한 아이템 발굴, 분석결과에 대한 해석, 공공기관의 자료 수

집 및 저장 등의 기능에 초점을 맞추는 것이 적합할 것이라고 함

- 따라서 초기에는 기초데이터 관리와 사업관리에 필요한 시스템으로 시작하고, 향후 이를 확장해 나가는 방향을 설정하는 것이 효과적일 것이라는 의견임

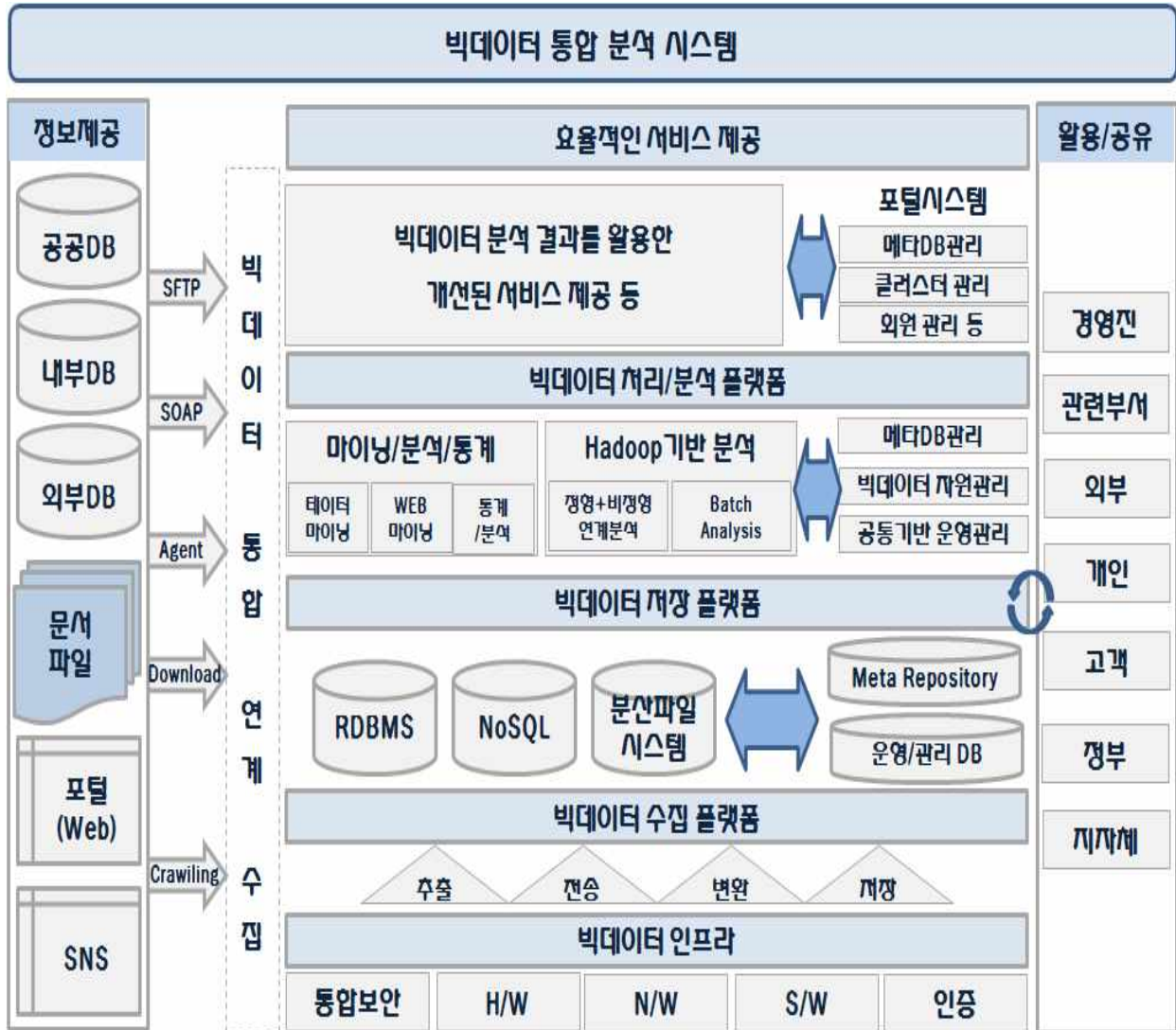
## □ 창원문성대학 빅데이터분석센터

- 창원문성대학에서는 빅데이터와 관련하여 분석 및 인력양성 사업에 초점을 맞추고 최근 이 분야에 특화된 전문가 5명을 유치했고, 빅데이터 분석을 위해서는 중소규모의 시스템을 구축하고 있는 상황이고, 대규모 분석이 필요한 경우에는 외부기관의 서비스를 활용할 계획에 있음
- 창원문성대학 빅데이터분석센터는 빅데이터 분석을 위한 시스템 구축은 장기적인 측면에서의 접근이 필수적이므로 초기에는 다음과 같은 분야에 초점을 맞춘 시스템 구축이 적합할 것이라는 의견임
  - 빅데이터의 효과적인 분석을 위해서는 기초 자료의 수집이 뒷받침되어야 하므로 초기에는 이러한 부분들에 초점을 맞춘 시스템 구축 필요
  - 빅데이터 분석을 고려하여 수집된 자료의 DB화
  - 초기시스템의 경우 PC Network 구축
  - 공공서비스 또는 공개서비스 활용 방안 마련

## 2) 시스템 구축 방향

- 빅데이터 사업의 효과적인 추진을 위해서는 전문인력과 함께 빅데이터 분석을 위한 데이터 수집, 솔루션 및 활용·공유를 위한 시스템이 구축되어야 할 것임
- 빅데이터 시스템은 크게 보안·H/W·S/W 등과 같은 인프라와 데이터의 추출·전송·변환 등의 기능을 담당할 데이터 수집 플랫폼, 빅데이터 분석을 위해 가공한 데이터의 저장 플랫폼, 처리 플랫폼, 서비스 시스템 등으로 구분할 수 있음

- 이와 같은 데이터의 수집, 저장, 가공, 분석 및 서비스를 고려한 일반적인 빅데이터 분석시스템 아키텍처는 그림25와 같음



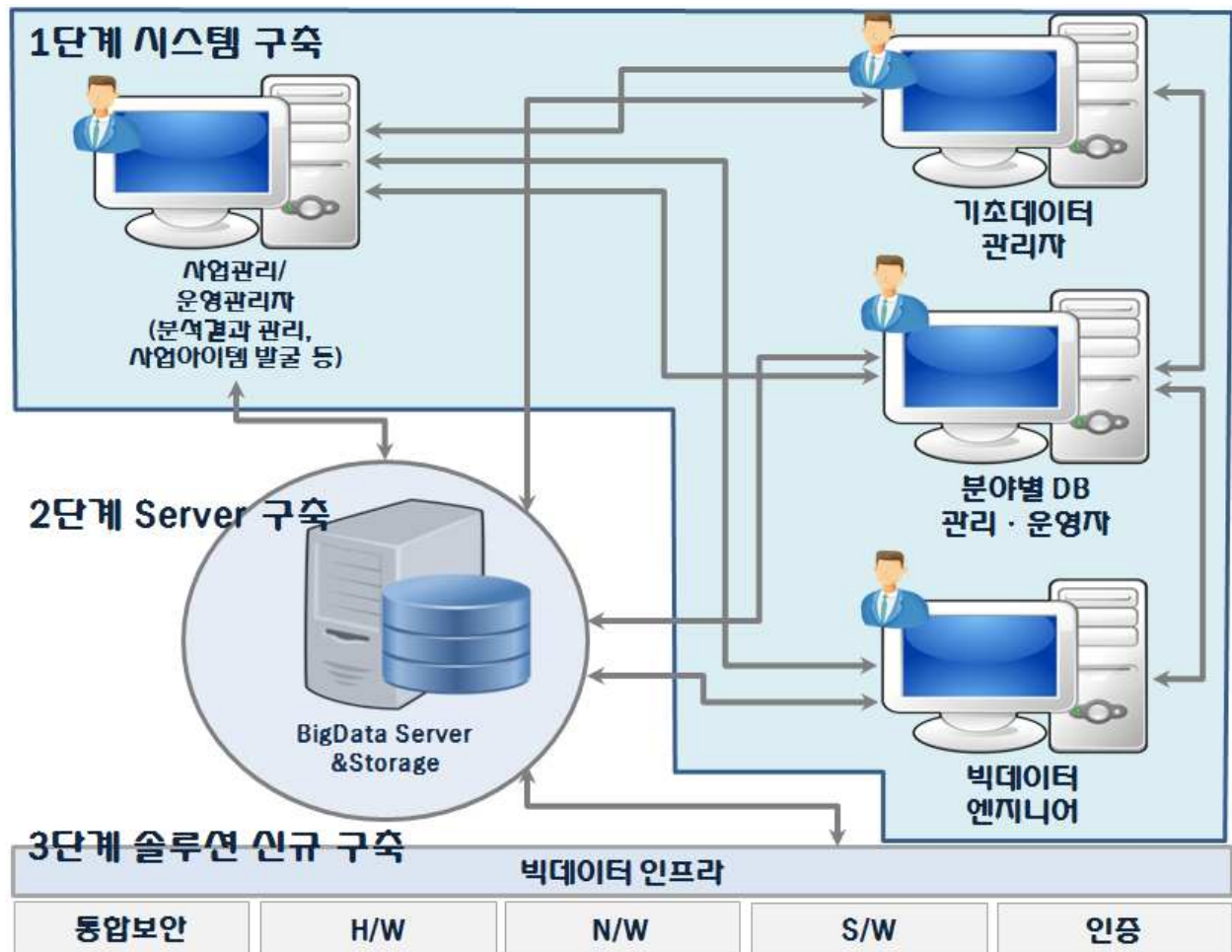
(그림25) 빅데이터 통합 분석시스템 아키텍처

- 그러나 빅데이터 분석과 관련한 경험이 없는 조직이 초기 단계에서부터 종합적인 시스템을 구축하는 것에는 무리가 있으므로 상황을 고려하여 단계적으로 구축하는 것이 적절할 것으로 판단됨
- 우선 빅데이터 분석 효율성 및 예산투입의 효과성 등을 고려할 때 크게 3단계 정도로 구분할 수 있는데 각 단계별 시스템 구축 방안을 정리하면 표24와 같음

(표24) 경상남도 빅데이터 분석시스템을 통한 단계별 추진 사항 및 소요예산

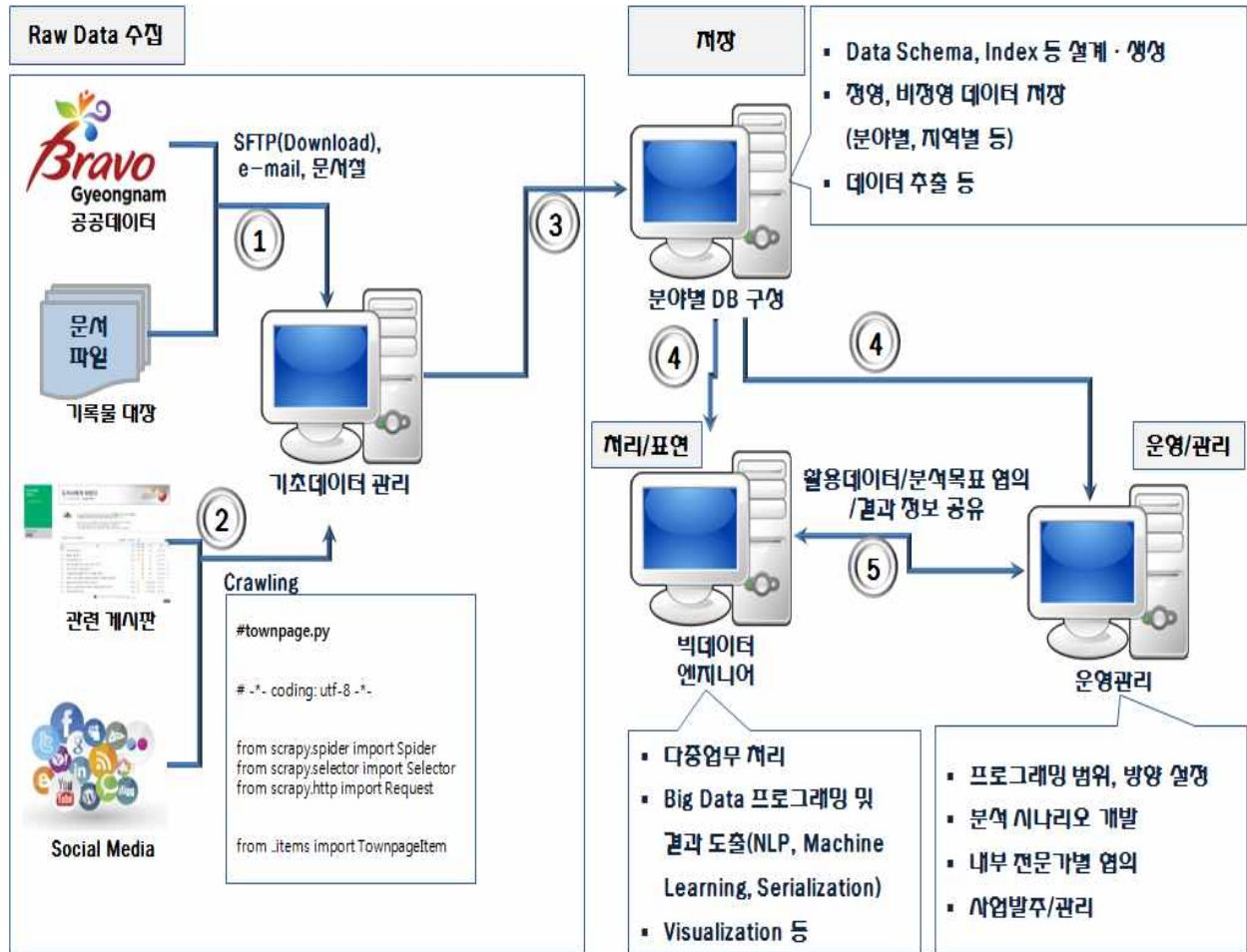
| 구 분 | 주요 수행내용  | 소요예산   |
|-----|--|--------|
| 1단계 | <ul style="list-style-type: none"> <li>○ 공공데이터(정형, 비정형)수집, 저장, 관리</li> <li>○ PC 수준에서 공공서비스를 활용하여 빅데이터 분석 실시</li> <li>○ 빅데이터 분석결과를 활용한 정책연구 수행 및 신규 사업아이템 발굴</li> <li>○ 시범사업의 경우 전문기관 위탁 및 지속적인 사업관리</li> </ul> | 5~6천만원 |
| 2단계 | <ul style="list-style-type: none"> <li>○ 포털사이트 및 소셜데이터 활용</li> <li>○ 데이터 분석 규모에 맞는 Server 및 Storage 구축</li> </ul>  | 3~4억원  |
| 3단계 | <ul style="list-style-type: none"> <li>○ 경상남도 전 분야 빅데이터 분석 전담</li> <li>○ 빅데이터 구축을 위한 Total Solution 구축</li> </ul>  | 8억원 내외 |

○ 각 단계별 주요 수행내용을 지원할 수 있는 시스템 구축(안)을 도식화하면 그림26과 같음



(그림26) 빅데이터 사업 추진을 위한 단계별 시스템 구축(안)

- 시범사업 추진을 위한 1단계 시스템의 경우는 PC 위주로 작업이 가능할 것으로 판단되고, 보다 세분화하면 Raw Data 수집, 저장, 처리 및 표현, 운영 및 관리를 위한 장비가 필요할 것으로 판단됨(그림27 참조)



(그림27) 빅데이터 사업 추진을 위한 1단계 시스템 구축(안)

- 먼저 Raw Data의 경우 경상남도 및 18개 시·군에서 보유하고 있는 정형·비정형 데이터, 관련 게시판, 포털사이트를 비롯한 소셜데이터를 말하고, 목적에 따라 기초데이터로 관리함
- 다음으로 Raw Data를 기초로 시범사업에 필요한 데이터를 다양한 형태로 분류하고, 분야별로 추출하여 빅데이터 분석의 기초자료로 활용함
- 1차 가공된 기초자료는 외부기관 제공 부문과 내부 처리 부문으로 구분 지

을 수 있는데 외부기관 제공 부문의 경우는 자료의 양이 방대하고, PC 단위에서 처리하기 어려운 데이터인 반면 내부 처리 부문의 경우 경상남도 현안과 관련한 사항으로 데이터의 양은 많지 않으나 PC 단위에서 신속하게 처리되어야 하는 것으로 정의할 수 있음

- 따라서 빅데이터 센터 1단계 시스템은 내부 처리 부문에 초점을 맞춘 PC 레벨의 시스템으로 구성하는 것이 적절할 것으로 판단됨
- 이외에도 분석, 처리, 표현, 운영 및 관리 등을 위한 기본적인 시스템이 구축되어야 할 것으로 사료됨
- 이상과 같은 시스템 구축을 통해 효과적인 업무를 수행하기 위한 각 부문별 기술적인 사항 및 전제조건 등을 정리하면 표25와 같음

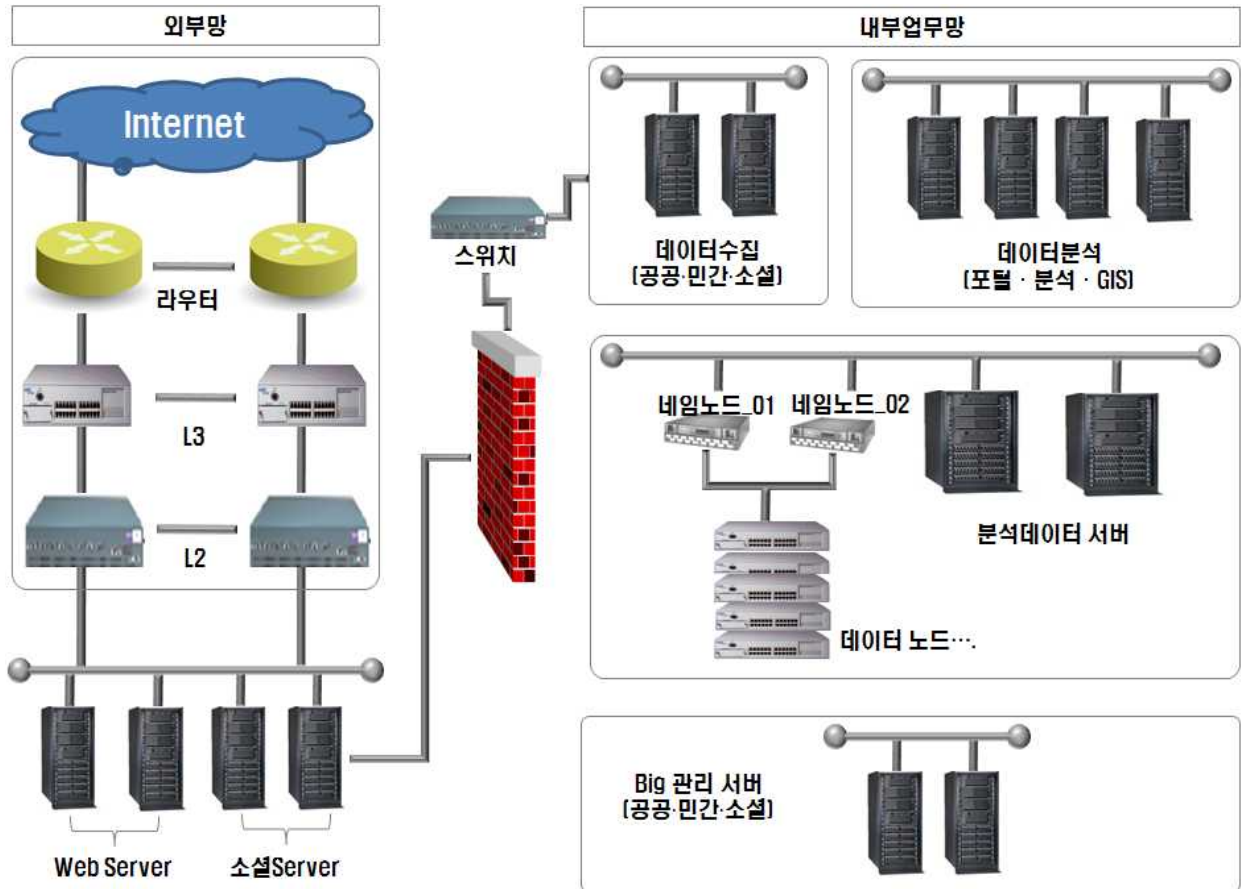
(표25) 경상남도 빅데이터 사업 추진을 위한 시스템 구축 시 부문별 주요 내용

| 구 분 | 주요 내용  |
|-----|--|
| 수집  | <ul style="list-style-type: none"> <li>○ SFTP : 관련 데이터베이스에 직접 연결하여 필요한 데이터 Downloading</li> <li>○ E-mail : 기존 자료(예: 엑셀, 한글 등) 수집</li> <li>○ 문서철(기록대장 등) 수집</li> <li>○ Crawling : 관련 게시판 및 소셜 데이터 수집</li> </ul> <p>⇒ 비정형 데이터의 경우 분석을 고려한 정리작업이 필요하고, SFTP 및 Crawling 등은 필요시 프로그래밍 작업이 수반되어야 함</p> |
| 저장  | <ul style="list-style-type: none"> <li>○ 데이터베이스(NoSQL 등) 모델링</li> <li>○ Index 등 작업</li> </ul>  |
| 처리  | <ul style="list-style-type: none"> <li>○ 자연어 처리</li> <li>○ 기계 학습을 통한 패턴 추출</li> <li>○ 데이터간 순서화</li> </ul>  |
| 표현  | <ul style="list-style-type: none"> <li>○ 데이터를 실세계처럼 구현(예: Crossfilter, Tangle, Modest Maps, Kartograph, Gephi 등)</li> </ul>  |
| 관리  | <ul style="list-style-type: none"> <li>○ 사업 및 시스템 전반적인 사항에 대한 이해도가 있어야 하고, 각 단계별 수준 검토</li> <li>○ 기초자료 수집에 따른 각 기관별 연계, 위탁사업 수행 시 분석 시나리오 개발/협의, 진행과정 점검, 분석 목표 선정</li> </ul>  |

## 7. 하드웨어 아키텍처

- 경상남도 빅데이터 사업 추진을 위해서는 다양한 형태의 하드웨어가 구축되어야 하는데 크게 데이터 수집을 위한 하드웨어, 여러 목적에 따른 분석을

위한 하드웨어, 분석결과를 저장할 수 있는 하드웨어, 서비스 제공 및 분석 데이터 축적을 위한 하드웨어 등으로 구분 지을 수 있음



(그림28) 빅데이터 사업 추진을 위한 하드웨어 아키텍처(안)

- 하드웨어는 라우터, 스위치, 서버, 스토리지, 보완서버 등으로 구성할 수 있고, 각 상황에 따른 하드웨어를 구축할 수 있을 것임
- 경상남도의 경우 단계별 사업 추진을 목표로 하고 있으므로 초기에는 기본적인 하드웨어 구성 즉, 저비용·고성능을 구현할 수 있는 오픈 아키텍처를 중심으로 한 하드웨어를 구축하고, 이를 기반으로 변화하는 상황에 신속하게 대처할 수 있는 분산네트워킹이 고려되어야 할 것임
- 그러나 관리 데이터의 증가, 여러 가지 요구사항에의 대응 등을 위해서는 장래 다양한 형태의 서버 및 스토리지 등을 구축하는 방향을 설정하는 것



이 적절할 것으로 판단됨

- 특히 하드웨어의 경우 고가의 장비들이므로 점차 확장하거나 통합할 수 있는 스케일 아웃 아키텍처가 기반이 되어야 할 것임

## 8. 소프트웨어 아키텍처

- 빅데이터 사업추진을 위한 전체적인 소프트웨어의 경우 외부망에서 크롤링한 정보를 수집, 분석, 서비스를 위한 소프트웨어가 필요하고, 이를 도식화하면 그림29와 같음



(그림29) 빅데이터 사업 추진을 위한 소프트웨어 아키텍처(안)

- 소프트웨어 아키텍처의 경우 하드웨어와 마찬가지로 수집·분석·서비스 등의 규모에 따라 달리 구성할 수 있고, 필요로 하는 기능에 따라서도 다양하게

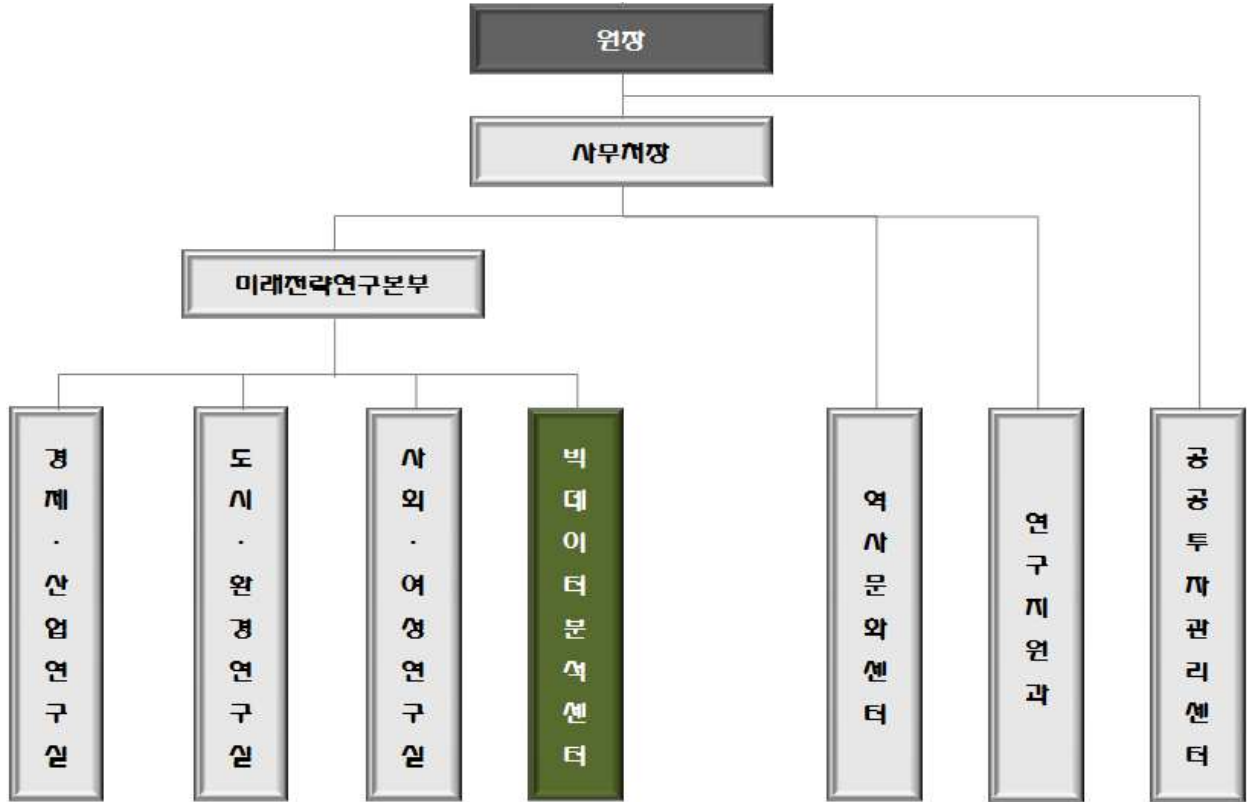
구성할 수 있음

- 경상남도 빅데이터 사업은 단계별 추진을 목표로 하고 있으므로 각 단계별 적정 소프트웨어의 활용이 요구되고, 장기적인 관점에서 범용성을 고려하여 구축하는 것이 타당할 것으로 판단됨

## 9. 추진 조직 및 사업 추진 절차

- 경상남도 빅데이터 사업의 효율적인 추진을 위해서는 몇 가지 전제되어야 할 사항이 있는데 정리하면 다음과 같음
  - 빅데이터 사업의 경우 장기 및 단기적 관점 모두에 대응이 가능해야 함
  - 다양한 분야에서의 정책 발굴 및 분야별 협업이 용이해야 할 것임
  - 정책 시행 기관과의 유기적인 연계가 가능한 조직이어야 함
- 이러한 사항들을 고려할 때 경상남도 정책연구 수행 기관인 경남발전연구원 은 빅데이터 사업을 효율적으로 추진할 수 있는 기본적인 체계가 갖추어져 있다고 판단됨
- 그러나 실제적인 사업의 추진을 위해서는 연구원 내 빅데이터 사업을 전담 할 수 있는 조직의 체계가 갖추어져야 할 것으로 사료됨
- 경남발전연구원내 빅데이터분석센터(가칭)는 미래전략연구본부 아래에 두고, 타 실과와의 효율적인 연계를 통해 빅데이터 관련 연구와 사업을 추진할 수 있도록 하는 것이 대안이 될 수 있음(그림 30참조)
- 빅데이터분석센터는 크게 사업기획, 자료수집 및 분석 등의 업무를 고려하여 센터장 및 2개팀(빅데이터 관리팀, 통계분석팀)으로 구성하여 운영하는 것이 효율적일 것으로 판단됨
- 센터장은 빅데이터분석센터 업무를 총괄(기획 및 분석)하고 빅데이터 관리 팀에는 사업관리, 데이터 수집 및 가공, 자료 관리를 위한 인력으로 구성하

고, 통계분석팀에는 모형개발, 통계분석, 조사관리를 위한 인력으로 구성함 (표26 참조)

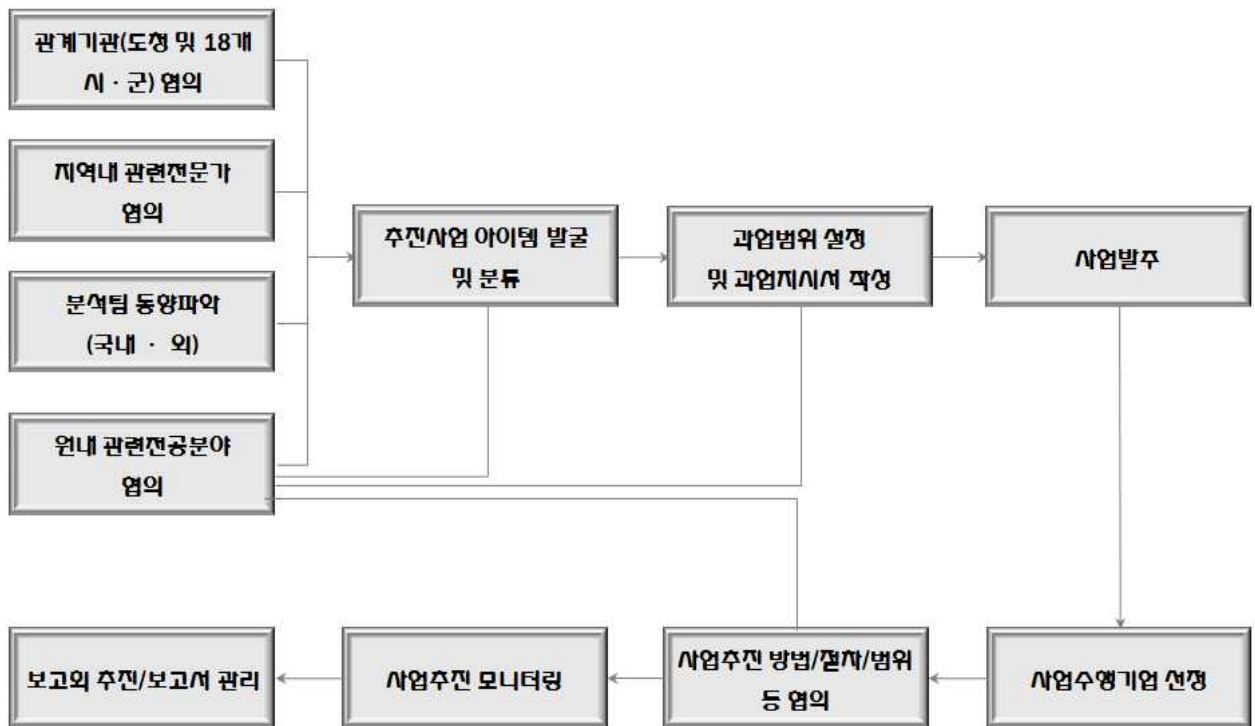


(그림30) 경남발전연구원 빅데이터분석센터 조직도

(표26) 경상남도 빅데이터분석센터 인력구성 및 담당업무

| 구분            | 전공            | 학력/직급                   | 인원 | 담당업무                                  |
|---------------|---------------|-------------------------|----|---------------------------------------|
| 사업총괄<br>(센터장) | IT, 최적화<br>관련 | 박사/<br>연구위원 이상          | 1  | ○빅데이터 분석센터 업무 총괄<br>○빅데이터 사업 기획 및 분석  |
| 빅데이터<br>관리팀   | 사업관리          | MIS<br>박사<br>(전문연구원)    | 1  | ○외주사업 관리/분야별 DB구축                     |
|               | 데이터수집·<br>가공  | 컴퓨터공학<br>석사<br>(전문연구원)  | 1  | ○빅데이터 자료 추출/분석 프로그래밍<br>○분석결과 시각화     |
|               | 자료관리          | 컴퓨터 관련<br>석사<br>(전문연구원) | 1  | ○자료 전산화/웹 자료 검색 등                     |
| 통계<br>분석팀     | 모형개발          | 통계학<br>박사<br>(연구지원직)    | 1  | ○통계모형 개발/조사계획 수립<br>○빅데이터 통계 분석 계획 수립 |
|               | 통계분석          | 통계학<br>석사<br>(전문연구원)    | 1  | ○경남경제동향/여민동락 조사분석<br>○빅데이터 관계 분석 등    |
|               | 조사관리          | 통계학<br>석사<br>(전문연구원)    | 1  | ○설문조사 분석/결과물 관리 등                     |

- 다음으로 빅데이터분석센터에서 실제 사업을 추진하기 위해서는 여러 단계의 과정이 필요한데 사업 추진 분야가 선정되면 사업을 통해 핵심적으로 발굴해야할 사항이 무엇인지, 가능한지 등의 파악을 위해 분야별 전문가들과 협의 및 관련 사례의 검토가 선행되어야 할 것임
- 이러한 사전 검토 과정을 거쳐 추진 사업 아이템을 세분화하고, 과업범위의 설정 및 사업발주의 과정을 거침
- 이후 사업수행기업이 선정되면 사업추진 방법/절차/범위 등을 협의하고, 사업 수행 전 과정을 모니터링하여 예외상황 대비 및 당초 목표로 한 결과물이 도출될 수 있도록 함



(그림31) 빅데이터 사업 추진 절차

# V 경상남도 빅데이터 시범 분석

## 1. 시범분석 개요

- 본 연구에서는 기 추진된바 있는 U-City 사업(양산시 경로당 복지시스템)을 통해 축적된 3년간(2012년~2014년)의 데이터를 기반으로 빅데이터 시범분석을 실시함
- 또한 노인들의 건강이 주변 환경과 연관성이 있는지 여부를 판단하기 위해 양산시 소재 산업단지(13개소) 또는 악취 측정 기록(11개소) 들을 함께 분석함

(표27) 경상남도 빅데이터 시범 분석 수집데이터

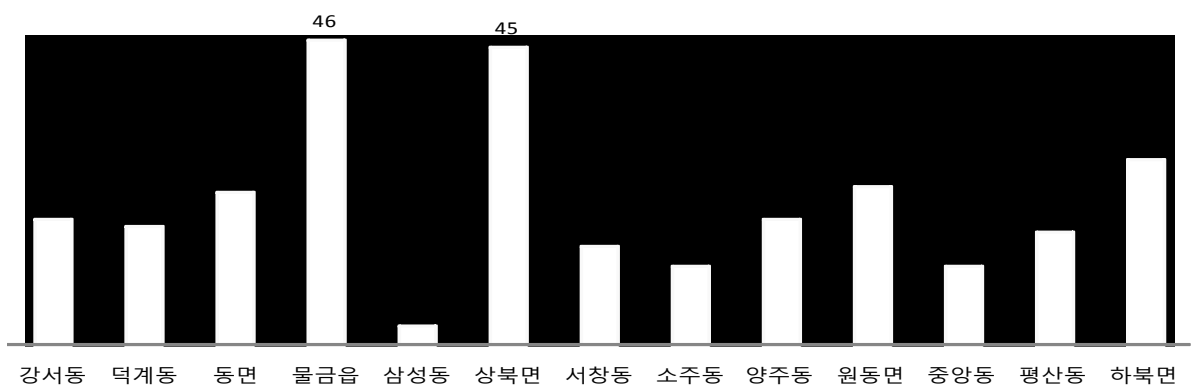
| 구분 | 획득데이터  | 비고                 |
|----|--|--------------------|
| 건강 | <ul style="list-style-type: none"> <li>○경로당 데이터 - 264 records</li> <li>○경로당별 인원(남여) 수 데이터</li> <li>○노인건강데이터_혈압 - 89,842 records</li> <li>○노인건강데이터_혈당 - 87,357 records</li> <li>○소변검사기록 - 8,334 records</li> <li>○노인별 건강 체크리스트 - 8,967 records</li> </ul> | 고혈압, 당뇨<br>고위험군 추출 |
| 산업 | <ul style="list-style-type: none"> <li>○양산시 산업단지 현황 - 13개소</li> <li>○산업단지별 업종 현황</li> </ul>  | -                  |
| 악취 | <ul style="list-style-type: none"> <li>○11개 지역 악취측정 결과 데이터</li> </ul>  | -                  |

- 양산시 경로당 복지시스템에는 경로당을 이용하는 노인들의 혈압, 혈당, 소변검사 기록 정보 등을 관리하고 있음
  - 혈압 데이터에는 지역, 경로당, 이름, 날짜, 혈압의 유무, 수축기 혈압, 이완기 혈압, 심박수, 판정범위 등에 대해 기록되어 있고, 총 89,842건이 기록되어 있음
  - 혈당 데이터에는 지역, 경로당, 이름 날짜, 혈당 유부, 혈당측정결과, 판정범위 등의 데이터가 87,375건 기록되어 있음

- 소변검사 데이터에는 유로빌리로젠, 요당, 빌리루빈, 케톤체, 요비중, 요잠혈, 요산도, 요단백, 요아질산염, 요백혈구 등의 데이터가 8,334건 기록되어 있음
- 기타 혈압, 당뇨 등의 질환 위험군에 대한 관리 데이터 8,967건이 기록되어 있음
- 이러한 데이터를 기반으로 혈압, 혈당 및 당뇨 등의 질환이 있는 노인들의 분포, 질환 종류에 따른 환자 수, 고위험군에 대한 관리 시 어떠한 부분이 부각되고 있는지 등에 대해 분석하고자 함
- 빅데이터 분석을 위한 데이터 가공은 Excel을 활용했고, 데이터 분석은 R i386 3.2.2 버전을 활용함

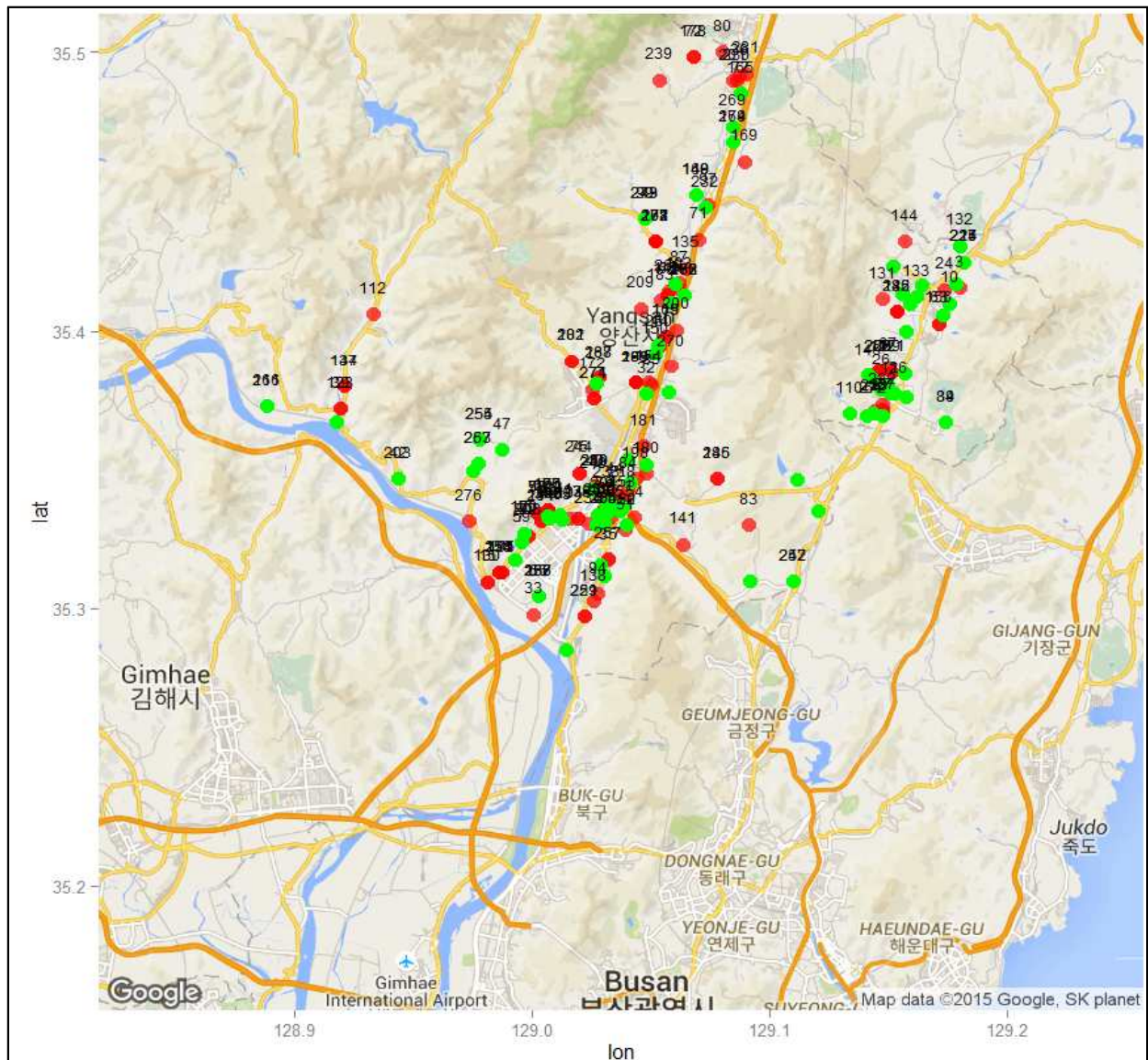
## 2. 분석결과

- 2010년 3월부터 현재까지 양산시 경로당 복지시스템에서 혈압, 혈당, 소변검사 등을 실시하는 노인인구 중 질환 위험군에 포함되어 있는 8,967건의 데이터 중 고혈압이나 당뇨병을 앓고 있는 환자 281명을 추출함
- 고혈압이나 당뇨병을 앓고 있는 환자 281명의 지역별 분포를 살펴보면 물금읍, 상북면이 각각 46명, 45명으로 질환자가 가장 많고, 다음으로 하북면, 원동면, 동면 등의 순으로 질환자가 많은 것으로 분석됨



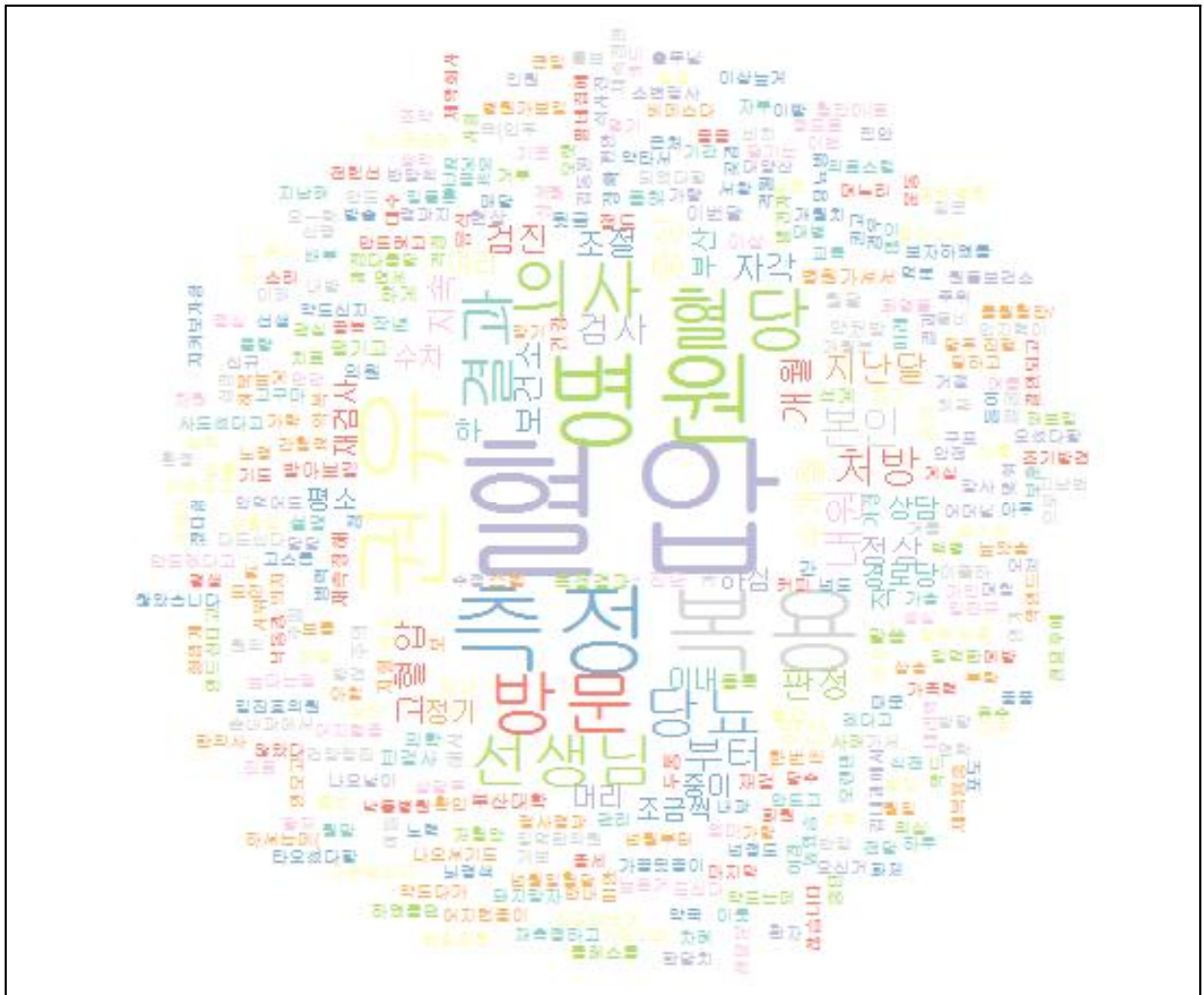
검사위치(경로당)의 위도와 경도 데이터를 파악하여 R프로그램에서 Google Map을 연동하여 표현하면 그림33과 같음

- 전체 281명 중 고혈압 환자는 189명이고, 당뇨병 환자는 92명으로 그림33에서 빨간색 점은 고혈압 환자이고, 초록색 점들은 당뇨병 환자들의 분포를 나타내고 있음
- 전반적으로 인구가 밀집해 있는 지역에 많은 질환자가 위치해 있는 것을 파악할 수 있음



(그림33) 양산시 노인 질환자 질환종류별 분포

- 이러한 질환자들에 대해 건강관리기록 시 언급된 단어들을 Wordcloud 분석으로 표현하면 그림34와 같음
- 질환자 중 고혈압 환자가 많으므로 혈압이라는 단어가 가장 많이 언급되고 있고, 병원, 측정, 권유, 혈당, 당뇨 등의 단어들이 많이 언급되고 있음을 파악할 수 있음
- 이러한 결과는 고위험군 환자들에게 병원진료를 권유하고 있으나 적극적인 치료가 이루어지지 않고 있음을 파악할 수 있음



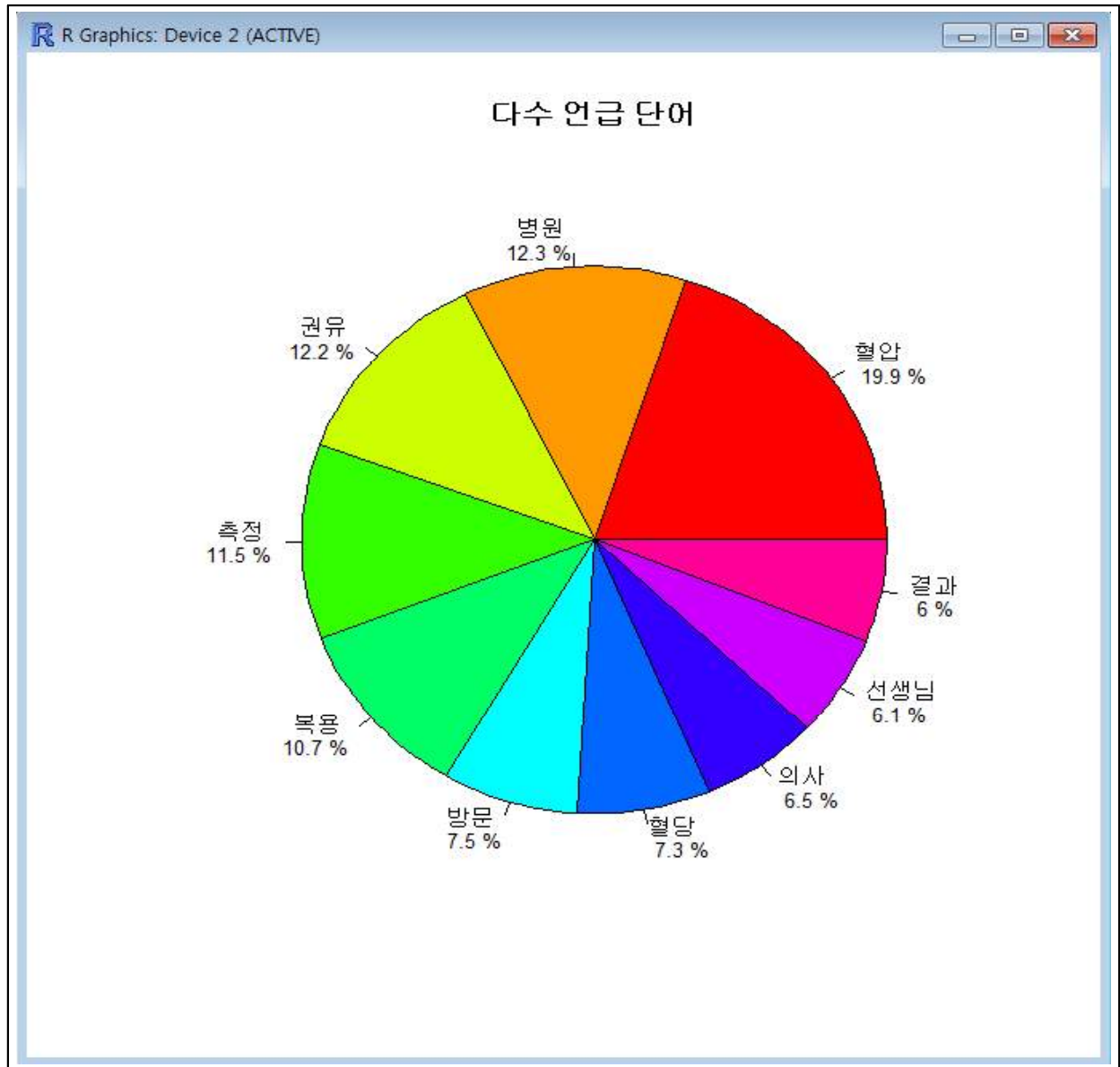
(그림34) 양산시 노인 질환자에 대한 체크리스트 분석

- 또한 노인질환자들이 의료시설을 이용하는 데에도 애로가 있는 것으로 미루



어 짐작할 수 있음

- 많이 언급된 단어들에 대해 보다 세부적으로 살펴보면 혈압이 약 20% 정도의 비중을 차지하며 가장 많이 언급되었고, 다음으로 병원(12.3%), 권유(12.2%) 등의 단어들이 많이 언급된 것으로 분석됨



(그림35) 양산시 노인 질환자에 대한 체크리스트에서 많이 언급된 단어들의 비중

- 추후 보다 세부적인 분석이 이루어진다면 분석결과를 기반으로 노인관련 의료시설 연계에 대한 방안도 마련할 수 있을 것으로 판단됨

### 3. 시사점

- 본 연구에서 기 추진된 U-City 사업 중 양산시 경로당 복지시스템에 축적되어 있는 노인건강정보를 기반으로 다양한 형태의 분석을 시행하고, 효과적인 결론을 도출하고자 했으나 한계가 있음
  - 관련 기관의 적극적인 협조 부족
  - 명확한 목표 설정 및 양산시 경로당 복지시스템과 직접적인 연관성을 가진 데이터(예: 의료시설 등) 확보 미흡
  - 관련 기관의 적극적인 협조가 있더라도 제공되는 데이터의 부정확성이 있어 이를 바로잡기 위한 많은 추가 작업 필요
  - 수집된 데이터의 경우 통계의 차원에서 정리는 가능하나 빅데이터의 관점에서 분석하기 위해서는 빅데이터 분석을 위한 기초자료 정리에 많은 시간 소요
  - 경로당 복지시스템을 통해 관리되는 정보가 혈압, 혈당, 소변검사 등으로 제한적이어서 지역 산업, 환경 등과 연관분석 실시 애로
- 이러한 과정을 통해 경상남도 빅데이터 시범사업 추진 시 기초데이터 수집, 분석 등의 단계에서 기본적으로 점검해야 할 사항들을 파악함
  - 필요 데이터 획득을 위해 어떤 기관과 연계해야하고, 지속적인 협조는 가능한지 파악
  - 빅데이터 분석을 고려한 기초데이터인지 파악
  - 기초데이터 보완 작업을 위해 필요한 사항(예: 인력, 시간, 예산 등) 파악
  - 획득된 데이터를 기반으로 분석했을 때 목표한 결과가 도출 가능한지 판단
- 향후 경상남도 빅데이터 시범사업 추진 시 이상의 점검사항들을 고려하여 추진한다면 보다 효과적인 결과를 도출할 수 있을 것으로 판단됨

## VI

## 결론

- 본 연구에서는 경남도의 주요 현안 문제에 경남도가 보유한 공공 데이터 등의 빅데이터를 활용하여 과학적이고 체계적인 행정을 구현하기 위해 경남도 주요 현안 중심으로 빅데이터 분석 사업을 발굴하고, 빅데이터 분석을 토대로 정책을 개발할 수 있도록 경남도 빅데이터 분석 및 활용 방향을 제시함
- 경상남도는 빅데이터와 관련한 경험이 부족하므로 목표로 빅데이터 활용체계 구축 및 정비, 빅데이터 전문성 확보, 빅데이터 플랫폼 구축 및 맞춤형 서비스 제공으로 설정하고, 이들 목표 달성을 통해 「Big Data 기반 도민 삶의 질 개선」이라는 비전 제시하였음
- 이를 위해 사업 추진 방향에 따른 기반인 빅데이터 플랫폼 구축, 전담조직, 하드웨어 및 소프트웨어, 기초데이터 수집·분석 체계 구축 등이 이루어져야 하기에, 향후 빅데이터 시범사업 완료 후 경남의 빅데이터 체계를 보다 확고히 하기 위해서는 연구원 내에 빅데이터 플랫폼 구축이 반드시 이루어져야 하고, 이를 통해 가공된 공공데이터 공유, 특정 분야 자료 실시간 공유, 빅데이터 분석 모델 확산 등이 이루어질 수 있을 것으로 판단됨
- 또한 21세기 원유로 비유되고 있는 빅데이터가 지역 경제 활성화에 도움이 될 수 있도록 빅데이터 스타트업 업체의 창업지원, 빅데이터 전문 인력 양성 지원사업도 반드시 빅데이터 사업에 포함되어 시행되어야 할 것으로 판단됨

## [참고문헌]

경기콘텐츠진흥원(빅파이 추진단), 빅파이추진단 중·장기 경영목표, 2015.1.28.

노규성, 사례분석을 통한 지방행정의 빅데이터 활용 전략, 2014.12. p.92.

미래창조과학부, 빅데이터 기반 미래예측 및 전력수립 지원계획, 2014.

\_\_\_\_\_, 창조경제 실현을 위한 K-ICT 전략, 2015.

한국정보화진흥원, 더 나은 미래를 위한 데이터 분석, 2013.4.

\_\_\_\_\_, 빅데이터로 진화하는 세상, 2012.5.

경기도청 홈페이지(www.gg.go.kr).

부산광역시 해운대구청 홈페이지(www.haeundae.go.kr).

서울특별시청 홈페이지(www.seoul.go.kr).

창원시청 홈페이지(www.changwon.go.kr).

경남신문, 빅데이터를 활용한 지방행정, 2015년 7월 23일자.

\_\_\_\_\_, 빅데이터를 활용한 지방행정, 2015년 7월 21일자.

\_\_\_\_\_, 창원시 빅데이터로 정책 효율성 높인다, 2015년 5월 21일자.