

서울형 빅데이터 공유·활용 플랫폼 고도화

민관 빅데이터의 융합·분석·활용으로 선제적 예측이 가능한 빅데이터 신행정을 구현하고 행정효율성 제고 및 시민의 정책신뢰 향상 도모

예산(안) 총괄

(단위 : 천원)

2014 최종예산 (A)	2015예산(안) (B)	증감 (B-A)	(B-A)*100/A
2,967,881	2,463,001	△ 504,880	△ 17%

사업설명

사업목적

- 데이터에 기반한 선제적 분석과 예측이 가능한 빅데이터 신행정 구현
- 민관 데이터의 융합·분석·활용으로 행정효율성 제고 및 시민의 정책신뢰 향상

사업근거

- 서울시 빅데이터 활용 기본계획(2013.4.20, 시장방침 제105호)
- 서울시 초협력 전자정부 종합계획(2013.5.29, 시장방침 제141호)
- 서울시 열린정부 3.0 종합계획(2013.6.14, 시장방침 제173호)
- 2015년 서울형 빅데이터 공유·활용 플랫폼 고도화 사업 계획
(2014. 7.31, 정보시스템담당관-9228)
- 빅데이터 기반 서울형 골목상권 분석서비스 구축 계획
(2014.10.15, 정보시스템담당관-12315)

사업내용

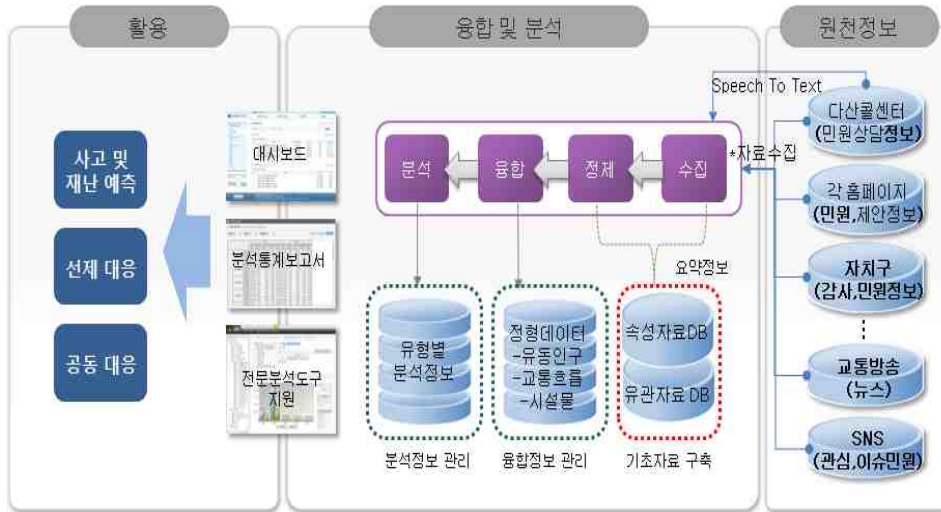
- 사업기간 : 2015. 1 ~ 2015.12
- '15년도 소요예산(안) : 2,463,001천원
- 사업내용

■ 사고 및 재난 사전 예보

- 필요성 : 행정 각 분야에서 사전에 예상하기 어려운 사건·사고를 미리 파악하여 선제 대응함으로써 시민 안전 확보

- 분석 및 모형 시나리오

- 120 다산콜센터에 접수된 시민 의견, SNS 상에서 분석된 시민 생각과 같은 비정형 데이터와 기존의 정형 데이터(유동 인구, 교통 흐름, 기후 상황 등)를 융합 분석하여 도시 안전의 위험 요소를 추출하여 사전 예보



- 활용데이터 : 다산콜센터 음성데이터, SNS 데이터, 홈페이지 게시판 데이터 등
- ※ 실시간 데이터 수집 및 분석과 활용을 목표로 2017년까지 단계별 추진

■ 교통사고 다발지역 분석(계속)

- 필요성

- 한국의 교통사고 사망자 수는 OECD 회원국중 최다수준(10만명당 10.5명 / OECD 평균 6.8명, '11년)
- 서울의 교통사고 사망자수는 주요 선진도시의 2~4배 수준

- 분석 및 모형 시나리오



- 활용 데이터 : 교통사고, 유동인구, 시설물, 도로시설물, 생각/감성

- 기대효과
 - 운전자 및 보행자 : 교통안전 교육을 통한 습관 개선, 법규 준수로 사고 예방과 감소
 - 서울시(도시교통본부): 사고 다발 지역 분석을 통하여 합리적인 교통안전 시설물 계획 수립

■ 자전거 시설 입지 추천

- 필요성
 - 자전거 도로 수요 대비 인프라 부족(일반 도로 대비 자전거 도로는 8.6% 수준, '13년)
 - 객관적인 데이터 부족으로 자전거 인프라 확장 지역 및 규모 판단 곤란
 - 특정 지역에 공공 자전거 편중 및 자전거 이용자 수의 증가에 따른 안전 및 도난 사고 증가
- 분석 및 모형 시나리오



- 활용 데이터 : 대중교통, 유동인구, 시설물, 대기환경, 생각/감성, 도로 현황 등
- 기대효과
 - 이용 시민 : 생활 교통수단으로서의 자전거 활성화, 편의 및 여가 생활 증대
 - 서울시(도시교통본부) : 과학적 데이터를 사용, 자전거 시설 필요 지역을 합리적으로 선정

■ 장애인 이동권 최적화 분석

- 필요성
 - 장애인 셔틀버스의 비합리적 이동 경로로 다중 교통 수단 환승 필요
 - 시각 장애인용 콜택시의 지역별/시간별 대기 시간 불균형

- 분석 및 모형 시나리오



- 활용 데이터 : 대중교통, 유동인구, 시설물, 생각/감성, 장애인 복지카드
- 기대효과
 - 장애인 : 차량 대기 시간 단축 및 이동 경로에 맞는 차량 이용으로 편의성 증진
 - 서울시(복지건강실): 수요 기반 장애인 차량 운영 효율화 방안 수립 및 약자 보호 시정 실현

■ 결핵환자 특성 요인 분석

- 필요성
 - 과거 후진국의 질병으로 여겨졌던 결핵이 급증하여 발병 및 사망률이 OECD 회원국 중 1위
 - 최근에는 노년층 뿐만 아니라 10대 중반부터 20대까지 발병률 급증 추세

- 분석 및 모형 시나리오



- 활용 데이터 : 인구분포, 시설물, 대기환경, 생각/감성, 환자정보, 진료정보

- 기대효과

- 고위험군 지역 시민 : 결핵 발병 패턴 분석을 통한 조기 예측 및 예방
- 서울시(복지건강실, 서북병원) : 서울시 결핵 예방을 위한 정책 및 신속 대응 체계 마련

■ 지역 축제 효과 분석

- 필요성

- 매년 지역별 축제는 증가하고 있으나 효과에 대한 객관적 근거 자료는 부족
- 축제 지역 거주민의 축제 만족도 및 문화 수요 충족도 판단 근거 자료 필요

- 분석 및 모형 시나리오



- 활용 데이터 : 인구분포, 교통흐름, 소비/매출, 시설물, 생각/감성

- 기대효과

- 축제지역 거주민 : 지역 문화축제 만족도 상승 및 문화 수요 충족
- 지역축제 주관부서/기관 : 지역 문화관광 축제 질적 향상과 축제지역주민과 공감대 형성

■ 소규모 자영업자를 위한 상권 분석(계속)

- 필요성

- 은퇴한 베이비부머 세대와 취업난을 겪고 있는 청년층 등 생계형 자영업자의 창업수 증가 추세로 밀집 포화상태의 자영업 시장
- 진입 장벽이 낮은 골목상권의 특성상 상권내 유사업종의 밀집 현상으로 단기 폐업 및 전환율이 높으므로 골목상권 내 서민 자영업자의 보호·육성 필요
- 상권내 공급, 경쟁구조, 소비자 트렌드 등 시장상황 변화에 적절히 대응하고 불확실성을 줄여 수익성 제고를 통한 상권의 안정과 활성화 필요

- 분석 및 모형 시나리오

분석 지역 골목 상권 특성 분석	모형 골목 상권 활성화 및 육성 모형																				
<p>업종별 자영업자 생존기간 추이</p> <table border="1"> <caption>업종별 자영업자 생존기간 추이 (단위: 년)</caption> <thead> <tr> <th>업종</th> <th>생존기간</th> </tr> </thead> <tbody> <tr> <td>부동산 판매업</td> <td>2.4</td> </tr> <tr> <td>학원</td> <td>3.0</td> </tr> <tr> <td>주점</td> <td>3.1</td> </tr> <tr> <td>음식점</td> <td>3.2</td> </tr> <tr> <td>현대 서비스업</td> <td>3.3</td> </tr> <tr> <td>숙박업</td> <td>3.7</td> </tr> <tr> <td>문구</td> <td>4.1</td> </tr> <tr> <td>차량</td> <td>4.4</td> </tr> <tr> <td>양육</td> <td>4.5</td> </tr> </tbody> </table> <p>평균 3.4년</p>	업종	생존기간	부동산 판매업	2.4	학원	3.0	주점	3.1	음식점	3.2	현대 서비스업	3.3	숙박업	3.7	문구	4.1	차량	4.4	양육	4.5	<p>거주인구밀도와 업종그룹 매출 상관관계</p> <p>상위 매출 그룹 (고밀도) 평균 매출 그룹 (중밀도) 하위 매출 그룹 (저밀도)</p>
업종	생존기간																				
부동산 판매업	2.4																				
학원	3.0																				
주점	3.1																				
음식점	3.2																				
현대 서비스업	3.3																				
숙박업	3.7																				
문구	4.1																				
차량	4.4																				
양육	4.5																				
<ul style="list-style-type: none"> •소상공인 매출/창업/폐업 분석 •블록별 세대구성/소득 구간 분석 •블록별 가구 소비 비율 분석 •상권별 이동 유입인구 패턴 분석 	<ul style="list-style-type: none"> •지역 상권별 목표 고객층 선정 모형 -1인/다인/청년/노인 가구 밀집도 •상권별 업종 회소성, 모방곤란성 추정 •상권별 전략 업종 선정 모형 																				

- 기대효과

- 자영업자 : 골목상권 지역에 대한 상권 분석정보의 생산과 제공으로 기존 자영업 자에게는 영업환경 개선과 매출증대 기회를 제공하고 창업 예정자에게는 합리적인 투자 의사결정 유도
- 전문컨설턴트 : 빅데이터 분석을 통해 자영업 환경 변화 등을 고려해 업종전환, 매출증대 방안, 마케팅 등 종합적이고 과학적인 컨설팅 수행
- 서울시(경제진흥실) : 골목상권의 영업 환경에 미치는 영향을 빅데이터 분석을 통해 파악하여 자영업자에 대한 시혜적 관점의 지원에서 정책적·거시적 관점의 지원으로 자영업자의 경쟁력 및 자생력 제고

■ 자치구와의 빅데이터 인프라 공동 활용 및 분석

- 필요성

- 최근 일부 자치구에서 빅데이터 활용의 중요성을 인식하여 자체적으로 시스템을 준비하기 위해 전문 인력과 시스템을 확보하려는 움직임이 있으나, 대부분의 자치구가 예산문제로 어려움을 겪고 있음

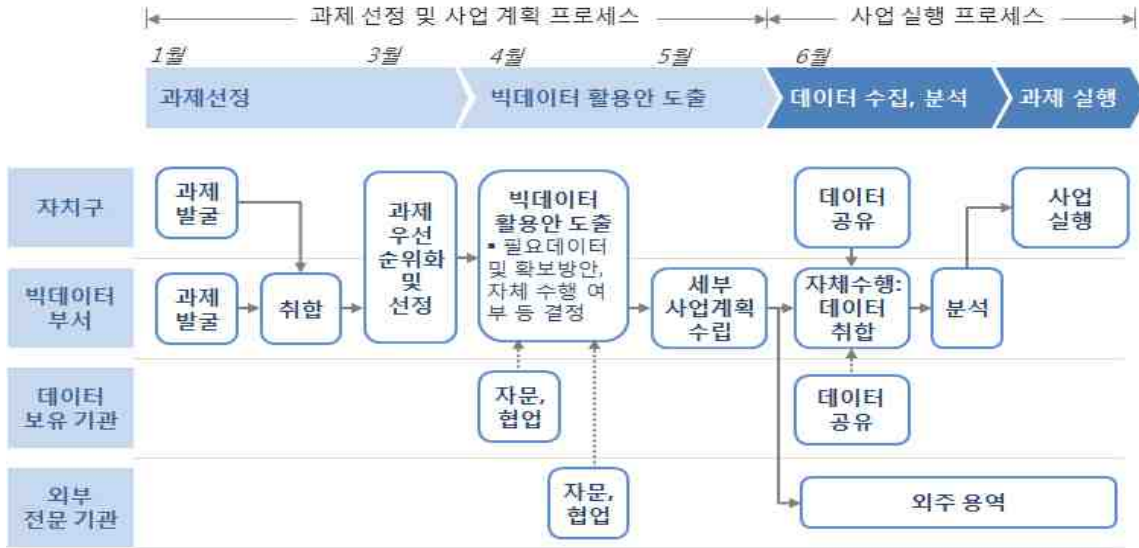
- 추진내용

- 서울시가 보유하고 있는 전문 인력과 분석 인프라를 활용, 일선 자치구에서 필요로 하는 빅데이터 분석 업무를 공동 수행

- 추진대상

- 가용한 자원(인력 및 인프라) 범위 내에서 협업을 통한 시와 자치구의 상승효과가 있는 과제
ex) 지역별 교통사고 상세 분석, 지역 축제 효과 분석 등

- 수행절차



- 기대효과

- 재정환경이 열악한 자치구도 데이터에 기반한 과학적 행정 수행 가능
- 유용한 데이터의 공동 활용과 참여자 다양화로 분석 결과 품질 상승 가능

□ 추진경위

- 서울시 빅데이터 활용 기본계획 수립(시장방침 제105호) : '13. 4
- 민관 융합 빅데이터를 활용한 심야버스 노선 최적화 : '13.3~7
 - 민간 이통사(KT)와의 MOU를 통한 30억건의 통화량데이터 활용하여 심야 유동인구 분석
 - 빅데이터를 활용한 도시문제 해결 방안 가능성을 확인한 혁신적인 성과 달성
- 빅데이터 활용한 종로 인생이모작지원센터 시설입지 분석 : '13.10~11
 - 유동인구, 거주지 근접도, 대중교통 접근성 분석하여 종로센터의 설치위치 타당성 검증
- 유동, 거주인구 분석을 통한 시정정보물 적정위치 분석 : '13.12~'14.3
 - 성별·연령별·소득별 인구 밀집지역 분석, 시정정보안내 매체현황 분석,
 - 시정정보 게시 최적지역 도출하여 시민 맞춤형 시정정보 제공
- 노인여가복지시설 수요 및 입지분석을 통한 중장기 정책수립 지원 : '13.12~'14.3
 - 어르신 이용현황 및 시설공급분석을 통한 노인여가복지시설의 수요공급분석
 - 자치구별 신규 시설 공급 우선순위 및 예상 후보지 도출
- 빅데이터 공유·활용 플랫폼 구축 사업 수행 : '14.7~'15.2
 - 택시이용 불편해소를 위한 택시 매치메이킹 서비스 개발
 - 교통사고 다발지역 분석을 통한 교통사고 감소 모형 개발
 - 장애인을 위한 효율적인 장애인 콜택시 운영 모형 개발
- 골목상권 이용 활성화 및 육성지원을 위한 상권분석 : '14.8~'15.7

□ 2015년도 예산(안) (단위 : 천원)

구 분	2013결산	2014예산(A)	2015예산(안)(B)	증감(B-A)
계	432,844	2,967,881	2,463,001	△504,880
사무관리비		-	40,150	40,150
공공운영비		-	42,711	42,711
전산개발비	284,344	2,412,532	2,052,140	△360,392
자산및물품취득비	148,500	555,349	328,000	△227,349
사무관리비		-	40,150	40,150
평가위원회 참석수당		-	3,150	3,150
전문가 자문회의 참석수당		-	7,000	7,000
빅데이터 활용 전략 컨퍼런스 개최			30,000	30,000
공공운영비		-	42,711	42,711
빅데이터 분석시스템 유지보수		-	42,711	42,711
전산개발비	284,344	2,412,532	2,052,140	△360,392
인프라 구축(S/W)		205,000	225,000	20,000
비정형 데이터 수집 및 가공		-	551,690	551,690
정형 데이터 수집 및 가공		-	196,530	196,530
분석 및 서비스 개발	284,344	2,207,532	756,270	△1,451,262
데이터 시각화		-	22,650	22,650
자치구 공동활용 및 분석		-	300,000	300,000
자산및물품취득비	148,500	555,349	328,000	△227,349
인프라 구축(H/W)	148,500	555,349	328,000	△227,349

□ 2015년도 예산(안) 산출 근거

- **비정형 데이터 수집 및 가공 : 551,690천원**
 - 대상 : 응답소, SNS(트위터, 페이스북 등), 전자문서 시스템
 - 내용 : 데이터 추출, 자연어 처리, 데이터 통합
 - 예산내역 : 인건비(551,690천원)
 - ▶ 데이터 추출(고급 3MM, 중급 16MM, 초급 18MM) : 291,760천원
 - ▶ 자연어 처리(고급 1MM, 중급 8MM, 초급 6MM) : 119,650천원
 - ▶ 데이터 통합(고급 1MM, 중급 8MM, 초급 9MM) : 140,280천원
- **정형 데이터 수집 및 가공 : 196,530천원**
 - 대상 : 상권분석 관련 데이터, 도시안전 관련 데이터, 복지시설 관련 데이터, 문화시설 관련 데이터, 교통 관련 데이터
 - 내용 : 데이터 추출, 정제/변형, 데이터 통합, 데이터 구입
 - 예산내역 : 데이터 구매비(120,000천원), 인건비(76,530천원)
 - ▶ 유동인구 데이터(연간구독 라이선스) : 120,000천원
 - ▶ 데이터 추출/변환/적재 (중급 3MM, 초급 1.5MM) : 35,780천원
 - ▶ 데이터 통합 (고급 1MM, 중급 2MM, 초급 2MM) : 40,750천원
- **분석 및 서비스 개발 : 756,270천원**
 - 대상 : 사고 및 재난 사전 예보, 교통사고 다발지역 분석, 자전거 시설 입지 추천, 장애인 이동권 최적화 분석, 결핵환자 특성 요인 분석, 지역 축제 효과 분석, 소규모 자영업자를 위한 상권 분석
 - 내용 : 다차원 데이터 분석, 모형 개발, 프로그램 개발, 서비스 운영
 - 예산내역 : 인건비(756,270천원)
 - ▶ 분석 및 서비스 개발 : 특급 10MM, 고급 23MM, 중급 26.5MM, 초급 22.5MM
- **데이터 시각화 : 22,650천원**
 - 내용 : 진행 과정 및 결과물의 차트, 인포그래픽, 모션그래픽 제작
 - 예산내역 : 인건비(22,650천원)
 - ▶ 인포그래픽(인건비) : 4,400천원, 모션그래픽(인건비) : 18,250천원
- **인프라 고도화 : 553,000천원 (HW :3.28억, SW:2.25억)**
 - 데이터 수집서버 2식 : 19,500천원 * 2 = 39,000천원
 - 스토리지 (NAS-30TB) : 140,000천원
 - 비정형수집기 SW(크롤링 툴 등) : 150,000천원
 - 비정형분석기 SW(NLP엔진 등) : 60,000천원
 - 비정형 분석서버 2식 : 20,000천원 * 2 = 40,000천원
 - 데이터 노드 증설 2식 : 20,000천원 * 2 = 40,000천원
 - 하둡솔루션 에이전트 SW 추가 : 2식 * 5,000 = 10,000천원
 - L2 스위치 2식 : 2,000천원 * 2 = 4,000천원
 - 서버 랙 1식 : 8,000천원
 - 백업 솔루션(SW) : 2대 * 2,500천원 = 5,000천원
 - 백업테이프 드라이브(30TB) : 57,000천원
- **사무관리비 : 40,150천원**
 - 제안서평가회 평가위원 참석수당 150천원*7명*3회 = 3,150천원
 - 전문가 자문회의 참석수당 100천원*7명*10회 = 7,000천원
 - 빅데이터 활용 전략 컨퍼런스 개최 : 30,000천원
- **공공운영비 : 42,711천원**
 - 빅데이터 분석플랫폼 유지보수 : 26,994천원
 - 주소정제툴 유지보수 : 3,055천원
 - ArcGIS for Server 유지보수 : 12,662천원

■ 자치구 공동활용 및 분석: 300,000천원

- 내용 : 서울시의 분석인프라를 활용하여 자치구에서 필요로 하는 빅데이터 분석 업무를 공동으로 수행
- 대상 : 지역별 교통사고 상세분석, 지역 축제효과 분석 등
- 예산내역 : 인건비(164,000천원), S/W 라이선스(136,000천원)
 - ▶ 인건비(고급 22MM, 중급 13MM, 초급 4MM) : 164,000천원
 - ▶ GIS Tool 라이선스 : 55,000천원
 - ▶ BI Tool 라이선스 : 81,000천원

□ 사업추진 절차

○ 집행절차

- 사업세부계획 수립 → 제안요청서 검토 및 계약심사 요청 → 입찰공고 → 협상에 의한 계약 → 사업시행 → 사업완료보고 → 사업비 집행

□ 사업효과

○ 연도별 사업추진 실적

2013년도	-서울시 ‘공공데이터’와 ㈜KT의 ‘트래픽데이터’융합을 위한 MOU체결 : ‘13.4 -유동인구 분석에 따른 심야버스 노선 수립 지원 : ‘13.3~7 · 심야버스 노선 타당성 검증 및 분석결과 일부 노선 변경에 반영 - 빅데이터를 활용한 종로 인생이모작 지원센터 시설입지 분석 : ‘13.10~11 · 유동인구와 시설이용자 현황 데이터 활용하여 종로센터의 접근편의성 분석 -노인여가복지시설 수요 및 입지분석을 통한 중장기 정책수립 지원 : ‘13.12~’14.3 · 노인복지시설 수요·공급 분석, 시설공급 우선순위 도출, 신규시설 적지 추천 -시정정보물 설치 적정위치 분석으로 시민 맞춤형 시정정보 제공 : ‘13.12~’14.3 · 성별·연령별·소득별 인구 밀집지역 분석, 홍보 매체현황 분석, 최적지역 도출 - 민·관 빅데이터 수집·정제 및 분석 빅데이터 플랫폼 구축 : ‘13.10~14.2 · 빅데이터 분석·공유가 가능한 플랫폼 구축으로 유동인구, 거주인구, 시설정보 등을 수집·공유하여 과학적· 선제적 행정서비스 제공기반 조성
2014년도	-빅데이터 기반의 교통서비스 모델 개발 타당성 검증 : ‘14.3~5 · 택시매치메이킹 서비스, 교통사고 위험지역 분석 모델 -서울형 빅데이터 공유·활용 플랫폼 구축 사업 : ‘14.7~’15.2 · 택시매치메이킹 서비스 개발, 교통 안전정책 수립지원, 장애인 콜택시 운영컨설팅 -빅데이터 기반 골목상권 분석서비스 구축 : ‘14.8~’15.7 · 골목상권 분석서비스, 골목상권 분석용 빅데이터셋 가공, 골목상권 빅데이터 플랫폼 확충

○ 향후 기대효과

- 빅데이터 분석을 통한 과학적이고 선제적 의사결정으로 미래 지향적 정책 변화
- 민관 데이터 융합 지식을 통한 다양한 공공 서비스 및 비즈니스 창출
- 사회·경제 정세의 실시간 분석으로 재난·복지·재정 등 미래 위험, 변화에 대한 선제적 예측 및 대응
- 수평적 데이터 활용을 통한 일하는 방식 혁신과 협업문화 조성
- 정확·정밀한 분석데이터 활용으로 행정효율성 제고 및 시민의 정책신뢰 향상
- SNS, 120민원 등의 빅데이터를 분석·활용하여 안전·시민복지·문화관광·경제 등 시정 제반 이슈에 대한 분석·예측을 제공
- 재난관련 민간·공공 데이터를 융합·분석하여 피해발생 위험요소 및 전조, 긴급이슈 등을 감지하여 피해·위험 적극 예방
- 시설물 적정규모 및 설치 최적지 도출로 공공 시설 이용대기시간 단축 등 대시민 편익 증대
- 다양한 형식의 대규모 빅데이터를 신속하게 처리하기 위한 빅데이터 인프라 고도화로 스마트 교통카드 데이터, 센서데이터 등 기존에 분석이 거의 불가능했던 빅데이터를 융합·분석함으로써 시정 전반의 다양한 정책 지원

□ 2015년도 추진방향 및 추진일정

○ 추진방향

- 빅데이터 분석을 통한 과학적·선제적 정책수립 및 시정현안 해결
- 공공과 민간의 데이터 수집·융합 및 분석을 위한 플랫폼 구축
- 주요 정책 의사 결정 지원을 위한 빅데이터 기반 행정 혁신 도모

○ 추진일정(단위 : 천원)

사업추진절차	추진기간	예산집행계획	추진세부내용
계		2,463,001	
사업세부계획 수립	2015.01.	-	2015년 사업계획수립
제안요청서 작성 및 사전협의 등	2015.02.~03.	-	제안요청서·계약심사 요청서 작성 및 검토
사업발주 및 사업자 선정	2015.03~04.	-	사업공고 및 계약체결
사업수행 및 준공	2015.05.~2015.12.	2,463,001	빅데이터 수집·가공·분석 및 서비스 개발

집행 결과 (단위 : 천원)

최근 3년간 이·전용, 이용·불용 실적

연도	당초예산 (A)	전년이월 (B)	이·전용 (C)	예산현액 (A+B+C)	집행	차년이월 (명사·사고)	불용
2011							
2012							
2013	433,400			433,400	432,844		556

주요 지적사항 및 조치결과

해당없음

실·국	부서명	담당자			
		정보기획관	통계데이터담당관	빅데이터전략팀장	고보영 (2133-4262)