

의안번호	제221호
보 고 년 월 일	2012. 7. 6. (제 11 회)

보 고 사 항
------------------

## 빅데이터 마스터플랜 추진현황 및 향후계획

제 출 자	국가정보화전략위원회 간사 맹형규 (행정안전부 장관)
제출년월일	2012. 6. 21.

## 1. 보고주문

- 『빅데이터 마스터플랜 추진현황 및 향후계획』을 별지와 같이 보고함

## 2. 보고이유

- 대용량 정보가 급증하는 빅데이터 시대에 자료를 체계적으로 분석하여 재난·환경문제 등에 사전 대응할 필요성이 제기되어, 정보화전략위원회에서 VIP께 보고(“빅데이터를 활용한 스마트 정부 구현”, ’11.11.7)
- 행정안전부에서 데이터 분석기반 정책수립·업무혁신, 맞춤형 대국민 서비스 제공 등을 위한 빅데이터 마스터플랜 수립 등 후속조치 추진

## 3. 추진현황

### 가. 빅데이터의 개념 및 범위

- (개념) 대용량 데이터를 활용·분석하여 가치 있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술
- (범위) 초기에는 기술 측면에서 접근하여 데이터만을 지칭하였으나, 수집, 저장, 검색, 공유, 분석, 시각화 등 관련 기술·도구 및 인재·조직으로 확대

## 나. 주요 사례

- (미국) 오바마 정부는 국방부 등 6개 연방기관이 빅데이터 기술개발에 2억 달러 이상을 투입하는 '빅데이터 연구개발 이니셔티브' 추진('12.3월)
  - 약 검색을 지원하는 국립보건원의 필박스(Pillbox)로 검색통계를 활용하여 주요 질병의 분포, 연도별 증가 등을 분석하고, 약의 효능 확인 비용(연간 5천만 달러) 절감효과 발생
- (싱가포르) 테러, 금융위기 등 국가 차원의 위험·기회요인을 선제적으로 대응하기 위한 국가위험관리시스템을 '04년부터 구축 운영
  - 해상테러 의사결정 지원, 조류독감 위협수준 평가·대응 방안 마련 등에 활용
- (권익위) 각 부처의 온라인 민원정보를 수집, 관리, 분석 및 예측하는 민원정보분석시스템을 '11년부터 구축 운영
  - 사회문제로 확산될 수 있는 민원의 사전 파악을 위하여 주요 이슈에 대한 국민의 의견 분석, 경보 및 예보 체계 구축
- (구글) 검색 통계를 이용해 오바마 정부의 노후 차량 보상 프로그램의 호응도를 분석하여 정부 예산이 부족할 것을 예측
  - 미국 정부는 새 차 구입 수요가 적을 것으로 예상하고 4개월 시행 목표로 10억 달러를 편성하였으나, 1주일만에 소진되고 20억 달러를 긴급 추가 편성
- (포스코) 가격 변동이 큰 철광석 등 자원을 적시 조달하기 위하여 고객의 수요, 해외 광산 상황, 런던 금속거래소의 광물 가격을 분석하여 철광석 가격을 예측하고 최적 구매시기 결정

## 다. 마스터플랜 추진 현황

- (T/F구성) 행안부, NIA의 업무 담당자로 빅데이터 마스터플랜 T/F를 구성하여 자문단 및 부처 회의 운영과 마스터플랜 집필 실무를 담당
  - T/F정보화전략실장 주재로 빅데이터 마스터플랜 목차(안) 토의(4.30)
- (자문단) 연구결과 발표 및 중간산출물 검토 등을 수행할 자문단을 학계, 공공, 산업계의 전문가로 구성
  - 빅데이터 마스터플랜 범위 및 주요 이슈사항에 대하여 전문가 자문회의 개최(5.11)
  - 빅데이터 분석 및 활용방안, 공공부문 적용 전략에 대하여 전문가 자문회의 개최(6.12)

※ NIA 주관으로 빅데이터 국가전략포럼 구성 및 회의 4회 개최
- (부처회의) 각 부처의 보유 데이터를 전수조사 중이며, 부처의 빅데이터 활용수요 조사 및 범정부적 공감대 확산을 위한 설명회 개최(6.15)

## 라. 모델 검증 추진 현황

- (배경) 인력에만 의존한 전조(前兆)정보 수집·관리 등으로 재난 전조업무 처리에 한계가 있으며, 집중호우에 의한 피해액 50% 이상이 내수침수에서 발생
- (모델검증) 서울지역 상습침수구역 대상 도시 내수침수 전조 감지·위험도 분석으로 빅데이터 활용 모델 검증

- 도시내수 침수 관련 정형정보(강우량, 관망자료, 지형자료, 침수이력, 배수시설 등) 연계, 뉴스, 트위터, 신고·제보 등 비정형정보 분석을 통한 위험도 분석 및 전조 감지 수행
- ※ 현재 모델검증 사업이 조달청 공고 중이며, 6월 중 계약 예정
- 구축 범위 및 주요 이슈사항에 대하여 소방방재청, 국립방재연구원 등 관련기관 회의 개최(6.7)

#### 4. 향후계획

- (마스터플랜) 6월까지 초안 작성, 7월까지 마스터플랜(안)을 작성하고, 8월에 마스터플랜 확정 및 국가정보화전략위원회 안건으로 VIP보고
- 마스터플랜(안)의 검토 및 보완을 위하여 부처회의, 자문단 회의 개최
- 각 부처 데이터 현황 조사결과를 정리하고, 도출된 과제의 실현가능성 검증 시 활용
- (모델 검증) 6월에 사업자를 선정하여, 8월까지 분석·설계 하고, 10월까지 개발 및 검증 완료
- VIP 보고(8월) 시 분석·설계 내용을 중심으로 보고

**【별 지】**

---

# **스마트 행정 구현을 위한 빅데이터 마스터플랜 현황과 추진 계획**

---

2012. 6.



**행정안전부**

## 1. 빅데이터란?

### □ 빅데이터 개념과 가치

- (정의) 대용량 데이터를 활용·분석하여 가치 있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술
  - 초기에는 기술 측면에서 접근하여 거대한 데이터 집합만을 지칭하였으나, 관련 기술 및 도구(수집, 저장, 검색, 공유, 분석, 시각화 등)로 확대

#### 《 빅데이터의 범위 》



- (가치) 최근 글로벌 컨설팅 기관 등은 빅데이터의 경제적 가치에 주목
  - IT의 주도권이 인프라, 기술, SW 등에서 데이터로 옮겨가고 있으며, 빅데이터를 IT 시장과 기술발전의 핵심 주제로 인식
  - 개인의 취향을 직접적으로 반영하는 소셜 데이터는 기업과 고객을 연결하는 쌍방향 소통 수단으로 활용되며 소셜 분석의 중요성 증대
  - 빅데이터에서 유용한 정보를 찾아내고 활용할 수 있는 기업이 시장을 선도하는 등 새로운 경제적 가치의 원천이 될 것으로 기대

## 2. 공공분야 필요성

### □ 국가 경쟁력 강화

- (동향) 주요 선진국에서 데이터 중심의 국가 발전전략 추진
  - (미국) 대통령 과학기술자문위원회는 ‘모든 연방 기관은 빅데이터 전략이 필요하다’고 강조하였으며(‘10.12월), 빅데이터 관련 연구 개발에 2억 달러 투입계획 발표(‘12.3월)
  - (영국) 데이터의 가치를 극대화하고, 경제·사회적 혜택 측면에서 데이터 공개와 접근 개선을 담당할 데이터 전략 위원회 신설(‘12.3월)
- (가치) EU는 공공분야 빅데이터 활용으로 비용절감, 부정·오류에 따른 손실 감소 및 세수가 증대되어 연간 1,500~3,000억 유로 규모 경제효과(연간 생산성 0.5% 증가) 예상
  - ※ 우리나라는 공공분야 빅데이터 활용으로 연간 10.7조원의 경제효과 예상

### □ 데이터 기반 스마트 전자정부 구현

- (현황) 주민정보, 종합소득정보 등 299개 기관 241종의 정보를 행정정보 공유시스템으로 연계하여 행정·공공기관의 업무처리에 활용
  - ※ 보건복지부는 국세청(종합소득), 국방부(군인퇴직연금) 등 12개 기관, 17개 정보를 연계하여 기초생활수급자 선정에 활용
- (한계) 정형자료를 기관 간 공동이용하여 업무 처리에 활용되고 있으나, 현재 시점의 데이터만 공유되며 연도별 현황 및 통계 등은 미제공
  - ※ 27개 기관 120종의 구비서류 정보를 2,988개 민원사무에 활용
- (목표) 정책보고서 등 지식정보, 부처 DB 등 각 부처가 보유한 정보를 수평 융합하고 데이터 분석에 기반한 의사결정으로 스마트 정부 구현
  - 범정부 차원에서 정책정보 등 데이터를 연계·융합하고, 가치있는 민간 데이터를 활용할 수 있는 인프라를 구축하고,
  - 데이터 분석을 통해 재난·복지 등에 적시성 있고 선제적인 데이터 기반 의사결정 및 정책 집행체계 마련



## II

## 국내외 활용 현황

### 1. 국내 공공분야

#### □ 국민권익위원회

- (민원분석) 권익위는 각 부처의 온라인 민원정보를 수집, 관리, 분류, 분석 및 예측하는 민원정보분석시스템을 '11년부터 구축·운영
  - 월별, 지역별 주요 민원 캘린더를 제작하여 사회적 이슈를 민원 지도 형태로 제공하고, 반복 민원에 대한 대책을 마련하였으며,
  - 사회문제로 확산될 수 있는 민원의 사전 파악을 위하여 주요 이슈에 대한 국민의 의견을 자동 분석, 경보 및 예보 체계 구축

#### □ 한국석유공사

- (오피넷) 석유공사는 고유가에 따른 소비자 부담을 줄이기 위하여 유가정보를 예측 제공하는 오피넷 시스템 구축
  - 국제 유가를 기반으로 1주일 뒤의 국내 정유사와 지역별 판매가격을 예측·자료로 제공하여, 유가에 민감한 국내 물가 안정에 기여
  - 국내 1,300여개 주유소의 현재 유가를 국민에게 제공하여, 인근 주유소 간 가격 경쟁체제 유도

#### □ 국립보건연구원

- (인체자원은행) 중앙인체자원은행을 운영하여 한국인의 질병 원인 규명, 진단·치료법 연구·개발에 필요한 인체자원 확보 및 제공
  - 전국 17개 병원에서 52만 명의 인체자원 정보\*를 확보하고, '11년까지 25만명의 정보를 454개 보건의료분야 연구개발(논문 159편)에 제공

\* 인체자원 정보(52만명) : 유전체 역학조사 등으로 확보한 일반인 정보 32만명, 병원에서 확보한 질환군 정보 20만명의 역학정보, 혈액, 조직, DNA 등 정보

## 2. 국내 민간분야

### □ 포스코

- (가격 예측) 가격 변동이 큰 철광석 등 자원을 적시에 조달하기 위하여 데이터 분석을 통하여 최적 구매 시기와 가격대를 결정
  - 고객의 수요 데이터, 남미·호주 광산의 상황, 런던 금속 거래소의 광물 가격 데이터를 분석하여 미래의 철광석 가격을 예측
- (생산 효율성) 생산 공정별 온도, 습도, 압력, 성분 등의 데이터와 불량률을 결합하여, 생산 효율성이 높아지도록 실시간으로 공정 제어

### □ SK텔레콤

- (상권분석) 지도와 연결된 유동인구, 업종별·월별 매출 등의 정보(10종)\*로 상권분석 서비스 제공
  - 자영업 창업 희망자가 업종별 매출 현황, 경쟁 매장, 잠재 수요 고객, 유동인구 등의 정보를 지도에서 직접 분석 가능
- \* 연결정보(10종) : 2,650만 SK텔레콤 가입자 동선(유동인구), 3,000만 OK 캐쉬백 회원 소비패턴, 현대카드 가맹점 결제, 부동산114의 상권 시세 등

### □ 다음소프트

- (소셜분석) 소셜 네트워크의 이슈와 관심 키워드의 실시간 모니터링으로 상황에 맞는 대응전략 및 마케팅 전략 수립 지원
  - 자연어 처리, 분석 기법 등을 이용하여 블로그와 트위터에서 형성되는 트렌드와 여론을 일반과 공공으로 구분\*하여 제공
- \* 일반 : insight.some.co.kr, 공공 : pub.some.co.kr

### 3. 해외 공공분야

#### □ 미국

- (연구개발) 오바마 정부는 국방부 등 6개 연방기관이 빅데이터 기술개발에 2억 달러 이상을 투입하는 '빅데이터 연구개발 이니셔티브' 추진('12.3월)
- (의료) 약 검색을 지원하는 국립보건원의 필박스(Pillbox) 서비스로 검색통계를 활용하여 주요 질병의 분포, 연도별 증가 등을 분석하고, 연간 5천만 달러 비용절감 효과 발생

#### □ 영국

- (공개) 영국은 공공부문의 정보 공유 및 활용을 위한 데이터 원스톱 서비스 제공(data.gov.uk)
  - (목표) 정부의 투명성 제고, 국민의 권리 향상, 데이터의 공개를 통한 경제적 사회적 가치 증대, 차세대 웹에서 주도권 획득 등
- ※ EU는 데이터 개방 전략을 발표('11년)하였으며, 우리나라, 미국, 호주 등도 공공정보 개방 서비스 운영 중

#### □ 싱가포르

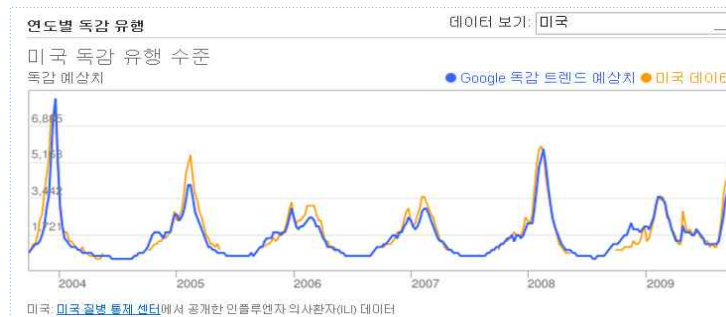
- (위험관리) 테러, 전염병, 금융위기 등 국가적 차원의 위험·기회 요인을 선제적으로 파악하고 대응하기 위한 국가위험관리시스템\*을 '04년부터 구축·운영
  - \* 국가위험관리시스템(RAHS) : Risk Assessment and Horizon Scanning
- (성과) 데이터 기반으로 정보를 조합하여 분석하는 기반을 확립하고, 해상테러 의사결정 지원, 조류독감 위협수준 평가·대응방안 마련 등 활용

## 4. 해외 민간분야

### □ 구글

- (독감 예측) 구글은 독감에 대한 인터넷 검색 통계를 이용하여 28개 국의 독감 유행 수준을 예측
  - (성과) 미국 질병통제예방센터\*보다 1주일 이상 앞서 독감 유행을 예측
  - \* 질병통제예방센터(CDC) : Center for Disease Control and Prevention

#### 《 구글 독감 예상치와 실측 데이터 》



- (차량 보상) 검색 통계를 이용해 오바마 정부의 노후 차량 보상 프로그램의 호응도를 분석하여 정부 예산이 부족할 것을 예측
  - 미국 정부는 새 차 구입의 수요가 적을 것으로 예상하고, 4개월 시행 목표로 10억 달러를 편성하였으나, 1주일만에 소진되고 20억 달러를 긴급 추가 편성

### □ 코노코필립스(ConocoPhillips)

- (장애 예측) 북해상에 있는 석유 시추선의 장애 예측으로 가동 중지로 인한 매출 감소 및 긴급 유지보수 비용을 예방
  - (내용) 버려지던 방대한 부품 데이터를 축적·분석하여 장애발생 징후 패턴을 발견하고, 필요 부품을 사전 공급·교체 및 가동중단 최소화
  - (성과) 가동중단 발생률을 80% 감소시켜, 생산성 20% 증가 및 운영비용 7억 달러를 절감

### Ⅲ

## 추진 현황

### 1. 마스터플랜 수립

#### □ 추진 체계

- (T/F) 행안부와 정보화진흥원(NIA)의 업무 담당자로 빅데이터 마스터플랜 T/F를 구성하여, 자문단 및 부처 회의를 운영하고, 빅데이터 마스터플랜 집필 실무를 담당
  - T/F 회의에서 빅데이터 마스터플랜 목차(안) 토의(4.30)
  - (자문단) 학계, 공공, 산업계의 빅데이터 관련 전문가로 자문단을 구성하고, 자문위원의 분야별 연구결과 발표, 중간산출물 검토 등 수행
    - 빅데이터 마스터플랜 범위 및 주요 이슈사항에 대하여 전문가 자문회의 개최(5.11)
    - 빅데이터 분석 및 활용방안, 공공부문 적용 전략 등에 대하여 전문가 자문회의 개최(6.12)
  - ※ NIA 주관으로 빅데이터 국가전략포럼 구성 및 회의 개최(4회)
  - (부처회의) 빅데이터 설명·공감대 확산을 위한 부처회의 등을 통하여 각 부처의 데이터 현황, 활용 수요 및 관련 정책 현황 조사
    - 빅데이터에서 연계·활용할 기초자료 조사를 위하여 각 부처에서 보유하고 있는 데이터의 현황 조사 중
    - 부처 실무자의 신속한 이해를 위하여 국내외 빅데이터 우수 사례에 대한 설명회를 개최(6.15)하고 활용 수요 조사
- (정책) 지경부, 방통위 등 각 부처에서 빅데이터 도입·활용 활성화를 위하여 수행하고 있는 정책 현황을 조사하여,
  - 전자정부 측면의 빅데이터 활성화 정책과 민간 기술·인력양성 측면의 빅데이터 연구개발 정책의 시너지 효과 극대화 도모

## 2. 파일럿 추진

- (배경) 재난전조(前兆) 정보의 수집·관리체계를 구축하고, 분석회의 개최·안전조치 시행 등 위험요인 사전 제거를 하고 있으나,
  - 인력에만 의존한 전조정보 수집·관리 등으로 전조업무 처리에 한계
  - 특히, 집중호우 피해가 급증하고 있으며, 피해액의 50% 이상이 내수침수에서 발생
- (파일럿) 서울지역 상습침수구역 대상 도시 내수침수 전조 감지·위험도 분석 파일럿으로 빅데이터 마스터플랜(안)을 검증

## IV 향후 계획

### □ 마스터플랜 수립

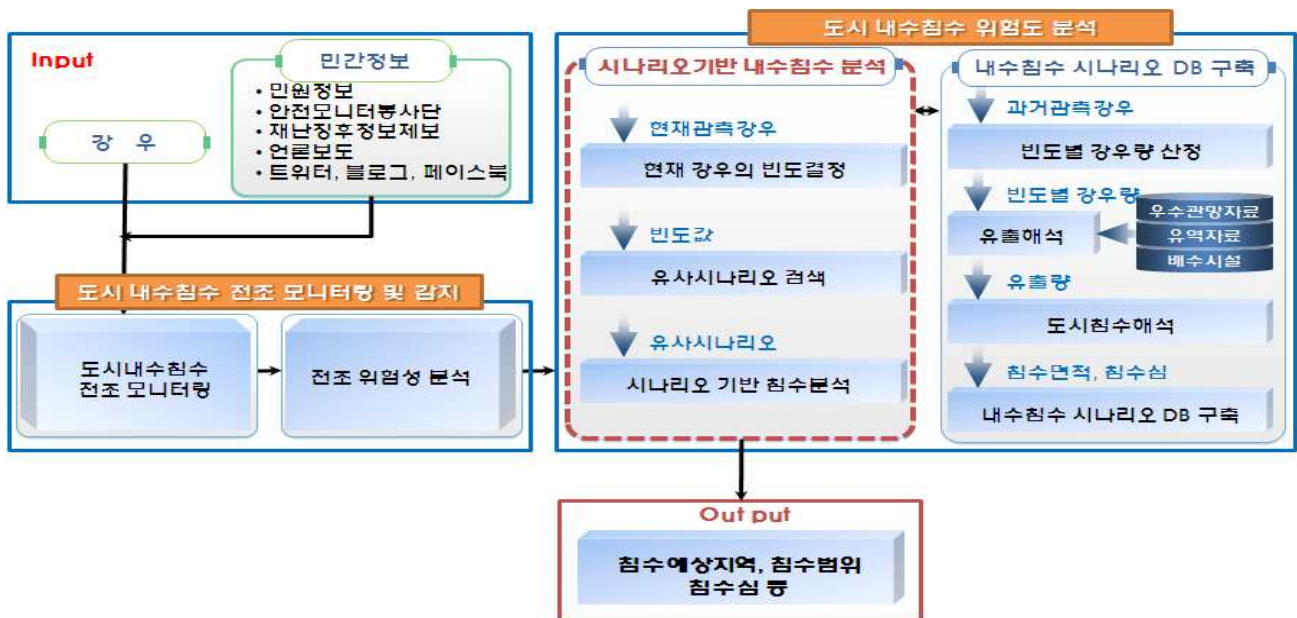
- (작성) T/F에서 담당 분야별로 빅데이터 마스터플랜(안) 작성
  - 초안 작성(6월) → 마스터플랜(안) 작성(7월) → VIP보고(8월)
  - ※ 정보화전략위원회의 VIP 보고 안건으로 8월말 보고 예정
  - 작성된 마스터플랜(안)의 검토·보완을 위하여 자문단 및 부처회의 지속 개최

### □ 파일럿 추진

- (과제 내용) 서울시 도시내수 침수 전조에 대한 분석예측용 파일럿 구축(국립방재연구원, 재난안전실, 소방방재청 공조)
  - (주요 내용) 도시내수 침수 관련 공공정보 연계, 뉴스 등 민간정보 수집체계를 구축하고, 모니터링을 통한 전조 감지 및 위험도 분석 수행

- (기대 효과) 하수관거의 사전 정비로 침수피해면적 저감(79.9%), 침수지역 홍수량 저감(88.4%) 예상
- (추진 일정) 9월까지 서울시 도시내수 침수 전조에 대한 분석예측용 파일럿을 구축하고,
  - 10월에 시험, 보완 및 빅데이터 활용 가능성 검증

### 《 빅데이터를 활용한 도시내수 침수 전조감지시스템 개념도 》



### □ 추진 일정

- (6월말) 마스터플랜 초안 작성 및 검토
- (7 ~ 8월) 마스터플랜 중간보고, 부처 및 자문단 회의 개최 등
- (8월말) 마스터플랜 확정 및 국가정보화전략위원회 안건으로 VIP보고
  - ※ 모델 검증은 모델 구축 기간을 고려하여 8월말 보고시 분석·설계 내용을 보고하고, 결과는 10월중 별도 보고

목 차	주요 내용
<b>제1장 빅데이터 시대와 스마트 정부</b>	
1.정보화 패러다임의 변화와 빅데이터	-정보화 패러다임의 변화: HW→SW →데이터 -새로운 패러다임 변화와 빅데이터
2.빅데이터 시대의 도래와 특징	-빅데이터의 개념과 가치 -빅데이터 시대의 특징과 새로운 가능성
3.공공부문의 빅데이터 범위와 의의	-공공부문 빅데이터의 개념과 범위 -공공부문 빅데이터 적용의 필요성
4.빅데이터 분석·활용 국내외 동향 및 시사점	-국내외 공공-민간 빅데이터 도입 현황 분석 및 시사점 -공공부문 데이터 보유 현황 및 수요
5.빅데이터 분석기반 스마트 정부 구현의 필요성	-스마트 정부의 특징과 빅데이터 활용 가치 -정부 의사결정의 품질 제고 및 선제적 행정의 필요성
<b>제2장 빅데이터 기반 스마트 정부 비전과 전략</b>	
1.비전과 목표	-빅데이터 기반 스마트 정부 비전 및 목표 수립
2.기본방향 및 추진전략	-비전과 목표 달성을 위한 추진전략 도출
<b>제3장 빅데이터 기반 스마트 정부 핵심과제</b>	
1.데이터 분석기반 정책결정 및 업무혁신 고도화	-정책 의사결정의 품질 제고와 일하는 방식의 혁신을 위한 빅데이터 활용방안 -유사업무 적용 확산
2.빅데이터 기반 선제적·맞춤형 대국민 서비스 발굴 및 제공	-맞춤형 생활공감서비스, 평생 질병예측 및 건강관리 프로그램, 선제적 위험예측, 국민안심서비스 등 대국민 서비스 발굴 및 제공 방안 -유사업무 적용 확산
3.데이터 기반 미래예측시스템 구축	-데이터 기반으로 중장기 정책방향 결정에 필요한 미래예측서비스 발굴 및 제공 방안



목 차	주요 내용
4.데이터 공유 및 분석 인프라 구축	-분석체계의 공통인프라(데이터 수집·저장 기반, 분석 공통모듈 등) 도출 -분석하는 부처에 필요한 인프라 도출
5.공공 데이터 관리체계 구축	-데이터 품질 진단·관리 -공공 데이터 수집, 저장, 가공, 처리 등 전 단계별 관리방안
<b>제4장 빅데이터 활성화 환경조성 과제</b>	
1.빅데이터 전문인력 양성	-데이터 사이언티스트 양성 과정 및 환경 조성 -산학 연계를 통한 빅데이터 전문인력 양성방안
2.정부·공공부문 데이터 분석역량 강화	-정부·공공부문 데이터 분석·활용 교육 프로그램 개발·운영 -데이터 분석역량 제고를 위한 교재 개발
3.빅데이터 분석 기술연구 지원	-빅데이터 저장·연계·처리·분석 기술 연구 -빅데이터 산업 육성 방안
4.빅데이터 시대 역기능 대응방안	-프라이버시 침해, 전자감시(빅브라더) 우려, 데이터 보안 등 부작용에 대한 대응 방안
5.빅데이터 분석·활용을 위한 법제도 정비	-빅데이터 기반 스마트 정부의 조기 정착을 위한 법제도 정비방안
<b>제5장 추진일정 및 기대효과</b>	
1.추진체계	-빅데이터 활용을 위한 범정부적 거버넌스 구성
2.추진 로드맵	-과제의 우선순위에 따라 로드맵 설계
3.재원조달 방안	
4.기대효과	