

데이터 산업 활성화 전략

- I-KOREA 4.0 데이터 분야 계획, I-DATA⁺ -

2018. 6.

관계부처 합동

순 서

I. 추진배경	1
II. 그간 정책의 한계	5
III. 비전 및 추진전략	9
IV. 추진과제	10
1. 데이터 이용제도 패러다임 전환	10
2. 데이터 가치사슬 초방위 혁신	14
3. 글로벌 데이터산업 육성기반 조성	22
V. 기대효과	25
VI. 추진체계	26
[붙임] 10대 산업 분야 빅데이터 구축 현황	28
[별첨] 데이터 산업 활성화 전략 [요약]	33

I . 추진배경

□ 데이터는 4차 산업혁명을 견인하는 핵심 동인

- 4차 산업혁명 시대 데이터가 모든 산업의 발전과 새로운 가치 창출의 촉매 역할을 하는 '데이터 경제(Data Economy)'로 패러다임 전환 중

* 세계 데이터 시장규모(IDC, '17): ('17) 1,508억 달러 → ('20) 2,100억 달러

- 4차 산업혁명의 진전에 따라 각종 IoT·센서 등에서 발생하는 대량의 데이터가 데이터 기반 산업·경제활성화를 견인하는 원동력으로 작용할 전망

* 세계 데이터량(IDC, '17): ('16) 16ZB → ('25) 163ZB, 10배 증가 전망

□ 데이터 기반 가치창출은 국가·기업의 혁신성장 수단

- 선진국은 국가 경제의 지속성장 및 일자리 창출을 위해 빅데이터 접목을 통한 주력산업의 재도약과 혁신성장을 도모

* (독일) 기계산업 '인더스트리 4.0', (일본) 로봇제조산업, (미국) ICT산업, (이스라엘) 항공·방위 산업

- 데이터가 기존 생산요소(자본, 노동)를 능가하는 경쟁원천으로 부상, 대규모 데이터를 보유하고 활용을 잘하는 기업이 시장 혁신을 주도

* (IBM) 약 3억 명 환자데이터 보유, 구글 딥마인드는 환자 100만 명 안구검사 기록 확보, (AirB&B) 8년간 축적한 데이터 분석, (알리페이) 5억명의 스마트폰 결제정보를 매초 2천건씩 축적

□ 데이터 경제 시대에 대응하는 데이터 산업육성 정책 필요

- 주요국은 미래 경쟁력을 좌우하는 데이터의 중요성을 인식, 데이터 산업 활성화를 위해 전략 수립과 투자확대 등 데이터 패권 경쟁 본격 돌입

* (美) 빅데이터 R&D 전략('16) / (EU) 데이터경제 육성 전략('17) / (日) Society 5.0 실현 데이터 활용 전개('17) / (中) 빅데이터산업 발전계획('17)

⇒ 데이터 경제의 급속한 진전에 대응, 혁신성장과 국민 삶의 질 향상을 동시에 실현하면서 새로운 기회로 활용하기 위한 전략 마련 필요

참고

데이터 경제 (Data Economy)

□ 개념

- 2011년 David Newman이 쓴 **Gartner 보고서**(How to Plan, Participate and Prosper in the Data Economy)에서 ‘데이터 경제’ 개념이 처음 등장

빅데이터, 오픈데이터, 연결데이터 등 데이터의 경제는 새로운 시대의 경쟁 우위를 주도하는 한 부분임을 의미하며, 선도 기업들은 데이터 경제의 단계를 이해하고 정보 공유를 통하여 정보 고립을 극복함

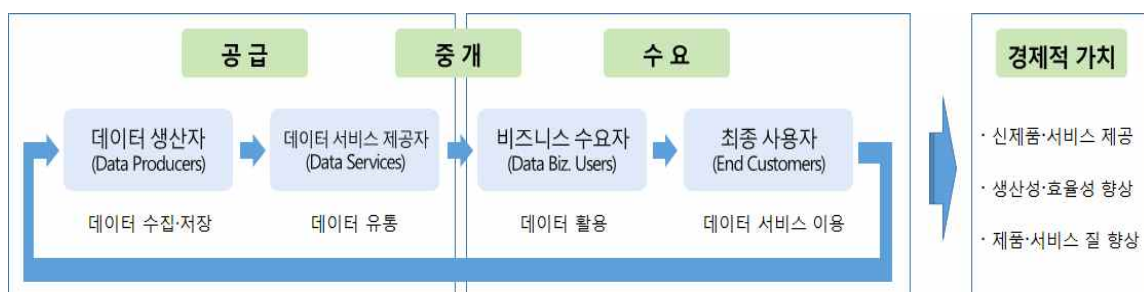
□ 확산 경과

- 2014년부터 유럽 집행위원회가 디지털 싱글 마켓 전략의 일환으로 ‘Data-driven Economy’ / ‘Data Economy’ 개념을 도입하면서 집중 조명
- 새로운 디지털 기회를 열고 디지털 경제의 글로벌 리더로서 유럽의 입지를 강화하기 위해 경제성장과 일자리의 동인으로 데이터 경제 표방
- 2016년 이후 IBM, MIT 등이 ‘Data Monetization’, ‘Data Capital’로 부연 전파

□ 가치창출 체계

- 데이터 생태계의 가치사슬(데이터의 수집·저장·유통·활용)을 기반으로 공급-중개-수요 시장을 통해 경제적 가치(신제품·서비스, 일자리 등) 창출
- ⇒ ①데이터 생산 → ②데이터 정제·가공을 통한 데이터·정보 유통 → ③기업, 정부, 공공기관 활용 → ④소비자, 시민 등이 데이터기반 혁신 서비스 이용

< 데이터 경제의 가치창출 체계 >



※ 출처 : Enter the Data Economy('17, EC), The Rise of the Data Economy('16, IBM) 재구성

참고

주요국가 데이터 경제 관련 정책

세계 주요국은 데이터 경제 선도를 위해 범부처 차원에서 주요 데이터의 확보, 산업·사회에 데이터 활용 확대, 데이터 분석 인재양성 등과 함께 안전한 데이터 활용제도 정비 등 종합적인 대책을 마련·추진

□ [미국] 빅데이터 R&D 전략 계획('16)

- (목표) 범부처 차원에서 빅데이터 7대 R&D 전략과 18개 세부과제를 제시하여 미래 빅데이터 환경 변화에 대응
- (주요내용) 빅데이터 기술 개발과 의사결정 도구 R&D 지원, 빅데이터 인프라 강화, 개인정보 보호와 윤리적 접근 추구, 빅데이터 인력 확충과 협력 생태계 구축

□ [EU] 데이터 경제 육성 전략('17)

- (목표) 유럽 내 통합 디지털 플랫폼(Digital European)을 기반으로 데이터 접근·분석·활용 강화를 통한 새로운 데이터 비즈니스 창출
- (주요내용) 데이터 접근권 강화, 기술 표준 제정, 법적 책임 명확화. 특히, 개인정보보호규정(GDPR) 제정으로 데이터 삭제권, 정보 이동권, 프로파일링에 대한 권리 등 개인정보 보호 강화와 합법적 데이터 유통 동시 추구

□ [일본] 미래투자전략-Society 5.0 실현을 위한 개혁('17)

- (목표) 5대 신성장 전략분야에 데이터를 활용한 기반구축
* 건강수명연장, 이동혁명실현, 공급망 첨단화, 쾌적한 도시만들기, 핀테크
- (주요내용) 새로운 사회의 인프라로서 '데이터기반(현실데이터 플랫폼)'을 구축하고 데이터 활용을 향한 제도정비, 교육·인재 역량강화, 혁신벤처 선순환 시스템 구축

□ [중국] 빅데이터산업 발전계획('17)

- (목표) 10개 이상 글로벌 빅데이터 선도기업, 500개 응용서비스 기업 육성
* 빅데이터를 중국의 인구·경제규모를 활용한 전 산업 발전 기회로 인식
- (주요내용) 데이터 개방 확대, 플랫폼·오픈소스 기술 지원, 빅데이터 전문 SW 수준 향상, 전문인재의 공급, 데이터거래소 등 생태계 조성

참고

데이터 시장규모 전망 및 주요 기업 사례

□ 데이터 시장규모 전망

- [세계] ('17년) 1,508억 달러 → ('20년) 2,100억 달러 (연평균 11.9% 성장)
 - '17년 지역별 시장 규모 : 미국(788억 달러), 서유럽(341억 달러), 아시아태평양(일본제외)(136억 달러) 지역 順
 - '17년 산업별 시장 규모 : बैंकिंग, 조립제조, 공정제조, 연방/중앙정부, 전문 서비스 분야 順 (5개 분야 724억 달러)
- [국내] ('17년) 6조 2,973억원 → ('20년) 7조 8,450억원 (연평균 7.6% 성장)
 - '17년 부문별 시장 규모 : 데이터구축/컨설팅(2조 9,291억원), 데이터 서비스 (1조 7,146억원), 데이터솔루션(1조 6,536억원) 順

※ [출처] Worldwide Semiannual Big Data and Analytics Spending Guide, IDC
2017. 4월 / 데이터 산업 현황 조사 보고서, 한국데이터진흥원 2017. 3월

□ 주요 기업 사례

- 세계 주요 기업들은 데이터 활용으로 제조 설비·과정 등을 **최적화**하거나 **생산량을 증가**시키고, 생산·유통·소비 전 과정에 **생산비용 절감**

GE	. 매출액의 75%가 자사 제품에 부착한 센서데이터를 통한 유지보수에서 발생 . 자사가 취급하는 선박항공기엔진·발전소터빈·의료기기 등에서 수집한 데이터 분석 결과를 고객에게 제공, 연간 200억 달러의 이익 창출을 예상
지멘스	. 제조 설비에서 발생한 데이터 분석, 생산라인을 재조정, 생산량 8배 증가
인텔	. 칩 제조 과정에서 발생하는 데이터를 분석해 제조비용 300만 달러 절감
코마츠	. 약 40만대의 건설장비 가동상황 데이터를 GPS 원격관리시스템으로 수집해 수요 예측과 수리, 중고차 가격결정 등에 활용해 10%가 넘는 영업이익률 제고
테스코	. 냉장 데이터를 분석해 영국·아일랜드 3,000개 점포에서 냉장비용 연 20% 절감

II. 그간 정책의 한계

- ◇ 글로벌 스탠다드에 뒤처진 데이터 이용제도, 양질의 데이터 부족, 취약한 산업·기술 기반 등이 4차 산업혁명 확산의 걸림돌로 작용
⇒ 4차 산업혁명 시대 데이터 경제를 선도하기 위한 대전환 시급

1 데이터 이용제도

- 우리나라는 세계적으로 엄격한 수준의 개인정보 규제*로 인해 데이터 활용이 크게 위축된 상황이며, 실제 보호수준도 높지 못한 현실

* 우리나라의 정보제공 동의제도 등의 규제는 OECD 주요국 중 가장 높은 수준(애널리시스 메이슨社, '14)

- 다양한 개인정보 유출사건들을 경험해 오면서 높은 수준의 규제체계가 확립된 반면, 개인정보의 활용에 대한 인식 및 법적 근거는 미흡

* 우리나라는 '14년 카드3사 개인정보 유출사태 이후 개인정보보호 규제가 한층 강화

- 개인정보가 포함된 빅데이터를 안전하게 활용할 수 있도록 비식별 조치 가이드라인('16)을 마련하였으나 활성화되지 못함

* [산업계] 비식별조치 절차가 지나치게 엄격·복잡하고 법적 면책이 보장되지 않음
[시민단체] 법적 근거가 없고 재식별 가능성 있음 ⇒ 가이드라인 활용기관 고발('17.11)

- 해외는 4차 산업혁명과 함께 새로운 ICT 기술에서 처리되는 개인정보가 증가함에 따라 개인정보의 보호와 활용의 조화를 이끄는 제도적 변화 시행

- EU는 GDPR*에 정보이동권을 신설하여 정보주체가 자기정보결정권을 행사할 수 있는 법적 근거 마련, 금융 등의 분야에서 Open API 형태 제공 본격화

* General Data Protection Regulation : 유럽내 모든 개인의 개인정보 데이터 보호규정('18.5.25 시행)

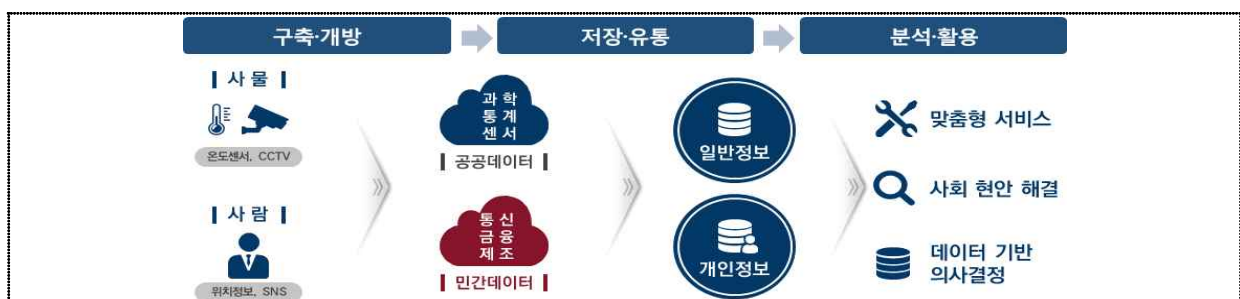
- 일본은 개인정보보호법에 제3자 제공에 동의를 요구하지 않는 익명화된 정보 유형을 도입하고 해당 정보의 안전한 유통을 촉진하는 규정 신설('15.9)

2

데이터 가치사슬

- 데이터의 구축·유통·활용 등 가치사슬 전반에 걸쳐 고품질 데이터 부족, 폐쇄적 유통구조, 산업·사회적 활용 저조 등 한계 노출
- (구축·개방) 4차 산업혁명 시대 新제품·서비스 개발에 필수인 자율차 스마트 시티 등 영역별 실제 데이터(Real Data)*와 AI 학습용 데이터 구축이 미흡
 - * (예) 자율주행차: 정밀지도 데이터, 도로주행 센서·카메라 데이터, 교통사고 데이터 등
- 또한, 지속적인 공공데이터 개방 확대에도 불구하고, 활용도가 높은 고부가가치 공공데이터는 양적으로 부족하고, 품질 역시 낮은 수준
 - * 한국 2.5만개, 미국 23.3만개, 영국 4.4만개('18.3) / 정확성 결여, 불일치, 중복·누락 등 발생
- (저장·유통) 빅데이터의 핵심 인프라인 클라우드의 이용률은 OECD 회원국 중 최하위권*으로, 조기확산을 위한 규제개선과 투자확대 시급
 - * 우리나라의 클라우드 이용률 12.9%로 OECD 33개국 중 27위('17)
- 개별기업 중심의 제한적 데이터 유통, 거래제도 미비* 등이 양질의 데이터 유통을 제한
 - * 데이터 거래시 애로사항(KDATA, '15): 불합리한 데이터 거래가격 33.9%, 거래절차/방법 미숙 11.9%
- (분석·활용) 데이터가 산업·사회 혁신의 촉매제로 역할이 기대되나, 산업적 활용은 아직 초기단계로서 산업 전반의 경쟁력 제고에 한계
 - * 국내 전체 사업체 빅데이터 이용률은 7.5%로 매우 낮음(NIA, '17)
- 국민의 삶과 직결된 교통·환경·안전이슈 등 사회문제 해결에 활용도 저조

< 데이터 가치사슬 생태계 >



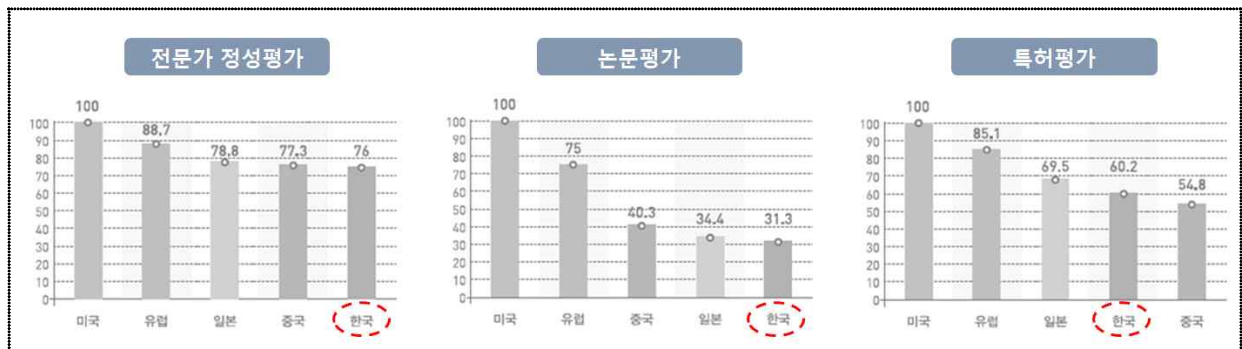
3

데이터 산업 육성기반

- 최근 빅데이터 R&D 투자 확대*에도 불구하고, 국내 기술수준은 세계 최고(미국) 대비 76%에 불과한 실정

* 빅데이터 관련 R&D 투자(억원): ('16) 126 → ('17) 155 → ('18) 176

< 빅데이터·클라우드 기술수준 평가결과('18.2, IITP) >



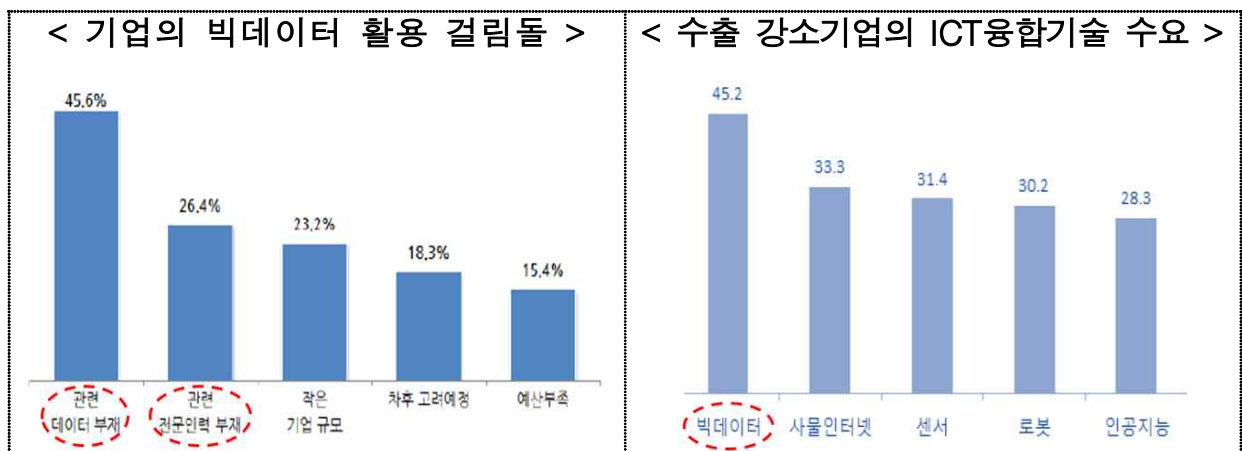
- 산업 전반의 빅데이터 활용 확산으로 빅데이터 분석 수요가 크게 늘고 있으나, 기업이 필요로 하는 전문인력*과 인프라**가 많이 부족

* 향후 3년간('18~'20) 빅데이터 분석가·개발자 등 전문인력 약 1.3만명 필요 전망(KDATA, '17)

** 빅데이터/AI 발전에 기반이 되는 10PF급 이상 슈퍼컴: 한국 0개 vs. 中 2개, 日 3개, 美 4개(과기부, '18)

- 국내 빅데이터 생태계에는 빅데이터 전문 기술을 보유한 스타트업이 부족하고, 글로벌 수준의 빅데이터 기술혁신 기업도 전무*

* 글로벌 100대 빅데이터 기술혁신 기업 중 국내 기업은 없음(美CRN社, '17)



※ 출처 : 한국데이터진흥원 보고서('18.3), 한국무역협회 보고서('17.12)

□ “데이터가 중요하다”

⇒ 다른 데이터와 결합이 가능한 데이터가 중요하다

- 어느 한 분야의 데이터만 분석하는 것보다는 이종 데이터를 결합·분석 시 예상치 못한 새로운 가치를 발견할 가능성이 높음

※ 농가 데이터 × 차량이동 데이터 → 조류독감 확산 예측

심야시간 통화 데이터 × 택시 승하차 데이터 → 심야버스 노선 수립

- 통계데이터, 정형데이터보다는 다른 데이터와 결합할 수 있는 원본 데이터의 활용 가치가 큼

□ “데이터가 없다”

⇒ 데이터 확보에 필요한 시간·비용을 투자할 여력이 부족하다

- 필요로 하는 데이터가 실제로 없을 수도 있음
- ICT 기술이 다양한 현장에서 활용되는 현실을 감안할 때,
 - ①데이터가 어디에 있는지 모르거나, 또는 ②데이터 확보에 많은 시간·비용이 소요되는 상황

□ “인공지능 시대, 데이터가 핵심이다”

⇒ 인공지능SW 개발을 위해 학습용 데이터가 필요하다

- 데이터는 오랜 옛날부터 분석·판단의 중요한 근거로 활용

※ 과학 분야에서 물체의 이동, 질량의 변화 등 관찰결과를 분석하여 이론 정립

- 인공지능 시대에 데이터를 주목하는 이유는 다양한 학습용 데이터를 기반으로 인공지능SW의 성능 향상

- 인공지능을 학습시킬 수 있는 데이터의 충분한 공급 필요

Ⅲ. 비전 및 추진전략

비전

데이터를 가장 안전하게 잘 쓰는 나라

기본방향

데이터 경제 활성화로
혁신성장과 삶의 질 향상

데이터 보호와 활용의 균형,
글로벌 스탠다드 지향



추진전략 및 과제

데이터 이용제도
패러다임 전환

데이터 가치사슬
순주기 혁신

글로벌 데이터산업
육성기반 조성

구축
개방

저장
유통

분석
활용

- 개인정보 자기통제·활용권 확립 (MyData 先도입)
- 데이터 안전활용 제도화 및 신기술 적용

- 산업별 실제 데이터·시데이터 전방위 구축
- 민·관 데이터 획기적 개방

- 공공·초산업 클라우드 이용 본격 확산
- 개방형 데이터 거래 체계 구축

- 데이터 기반 산업경쟁력 제고
- 빅데이터 활용 사회문제 해결 강화

- 빅데이터 선도기술 확보 (선진국 대비 90%↑ 수준)
- 미래수요 대응 전문인력 양성 (5만명 확충)
- 빅데이터 전문기업 성장 지원 (강소기업 100개 육성)

IV. 추진 과제

[전략 1] 데이터 이용제도 패러다임 전환

- ◇ 글로벌 수준에 부합하도록 ▲ 정보주체인 개인이 스스로 통제·활용하는 패러다임으로 전환, ▲ 개인정보의 안전한 활용 전방위 지원(법제화·안심존 구축·신기술 적용·EU GDPR대응) 추진

1 데이터 이동권 확립 : 국민의 데이터 주권 찾기(MyData)

- 정보주체가 기관으로부터 자기 정보를 직접 내려받아 이용하거나 제3자 제공을 허용하는 방식으로 데이터 활용체계를 정보주체 중심으로 전환(예: 美 블루버튼)
- ⇒ (MyData 선도입) 개인정보 관련 법 개정 없이도 시행 가능하고, 비식별조치 시 데이터 활용가치가 낮아지는 한계도 극복 가능

< MyData 개요 >



- (대규모 시범사업) 기관이 보유한 나에 관한 정보(예: 건강검진 결과) 및 내가 제공한 개인정보(예: 카드 사용내역, 통신 사용량)를 활용하는 시범사업을 대규모로 추진('18, 금융 등 2개 → '22, 10개 분야 이상 확대)

< 주요 분야 시범사업(안) >

분 야	내 용	참 여
[의료] 건강관리	건강검진 결과가 스마트폰 헬스앱으로 다운로드, 걸음수·심박수 등 각종 건강정보와 통합 관리 ⇨ 실시간 건강관리	병원, 휴대폰 제조사 (국대 5대 종합병원 확대)
[금융] 자산관리	계좌거래, 카드구매 내역 등을 Open API 형태로 제공받아 보다 빠르게 자산 통합조회 및 맞춤형 상품추천 ⇨ 안정적 재테크	핀테크업체, 은행·카드사 (100만 고객 대상)
[통신] 요금제추천	통신사가 보유한 이용자의 음성 및 데이터 사용량을 다운로드 받아 최적의 맞춤형 요금제 추천 ⇨ 가계 통신비 절감	KAIT, 통신사 (20만 고객 대상)

- (민간참여 확산) 데이터 보유기업에 대해 국민 편익제고 차원에서 MyData 자발적 참여를 권고하되, 공공성 높은 분야는 제도로 정착

참고

해외 주요국 마이데이터 사례

□ [미국] '11년 정부 주도로 '열린 정부를 위한 실행계획'의 추진 과제로 스마트공시(Smart disclosure) 추진

- 의료(블루버튼), 에너지(그린버튼), 교육(마이스튜던트버튼), 태양광(오렌지버튼) 분야에서 개인데이터의 다운로드 서비스 제공
- 특히, 의료분야는 '10년 재향군인들을 대상으로 의료정보 다운로드 서비스를 제공, '12년부터 미국 내 모든 환자를 대상으로 확대 적용*

* 약 3백만 명이 의료정보 다운로드 경험이 있으며, 약 16,000개 의료기관이 블루버튼 기능을 지원('16년), 애플 헬스앱 연계·활용('18.1월)

※ 애플 헬스앱은 ①의료기관의 의료정보를 다운로드 받아 저장하고, ②저장된 의료정보를 다른 앱과 공유 활용('18.1월), 現 39개 의료기관이 참여하며 지속 확대 전망

< 블루버튼을 활용한 민간 건강관리 서비스 >



□ [영국] '11년 정부 주도로 'Better Choice, Better Deals' 정책의 추진 과제로 midata 추진

- 금융, 에너지, 모바일 분야에서 소비자 데이터 제공을 의무화하고 제공 방식 및 전담기관을 지정(기업규제개혁법, '13.4월)

□ [스웨덴] 의료공공기관*(Inera)이 의료정보 포털을 만들고, 이를 통해 개인이 의료정보에 직접 접근할 수 있는 서비스를 제공

* 스웨덴의 21개 주 의회 및 지방정부 공동 소유(약 300명 근무), 주요서비스는 ①전화로 헬스케어 상담(월 45만 건), ②의료정보 온라인 제공, ③환자 분석정보 제공, ④병원 예약관리

- 스웨덴 인구 약 1천만 명 중 41백만 명이 의료정보 포털에 가입, 이중 1.3백만 명이 의료정보 온라인 제공 서비스를 이용('17.6월)

※ 20개 주에서 의료정보 온라인 접근 가능, 마지막 1개 주는 현재 구현 중

2

개인정보 안전한 활용 촉진

- (개인정보 제도화) 사회적 합의를 거쳐 가명·익명정보 개념을 명확히 하고, 위치정보에서 사물위치정보 제외 등 법 개정 추진(개인정보보호법·위치정보법 개정 '18)

* 시민단체·산업계·정부가 다함께 참여한 개인정보 해커톤 개최 등을 통해 사회적 합의 도출

< 제도 개편 방향(안) >

- (개인정보 범위 명확화) 개인정보보호법상 개인정보의 범위를 합리화하고, 위치정보법상 위치정보에서 드론 등 사물위치정보를 제외
- (비식별 조치 법제화) 비식별 조치 법적 근거를 명확히 하여 데이터 활용 촉진
 - * 개인정보 법적 개념체계를 개인정보 가명정보(연구, 통계 목적 활용), 익명정보(법적용 제외대상)로 구분
- (사전동의 제도 개선) 개인정보 수집·이용 및 제3자 제공시 서비스 개선 등 합리적 이유라고 판단될 경우 사전동의 예외 검토

- (데이터 안심존 구축) 데이터 자체의 반출은 안되고 데이터 분석 및 AI개발 결과만 반출하는, 보안환경을 갖춘 데이터 안심존* 제공('19~)

* 이용자가 물리적 보안 공간 또는 보안네트워크 원격시스템을 통해 다양한 분석SW(오픈소스툴 등)를 활용, 저장된 데이터(표본DB/맞춤형DB)를 분석해 볼 수 있는 온·오프라인 샌드박스 개념 형태

- 개인정보 비식별조치 기술에 대한 대국민 신뢰 제고 및 저변확대를 위해 국민과의 소통을 다양하게 추진

- (개인정보 비식별조치 콘테스트) 비식별처리 데이터의 유용성을 검증하고 재식별 시도·공격을 겨루는 콘테스트 개최(KISA, '18~)

* ('18년) 대회규칙 등 마련(~6월), 참가 팀(대학, 기업 등) 모집 및 약 10개팀 선발(~8월), 전문교육 실시 후 1차 예선대회(10월), 2차 본선대회(11월)

- (비식별조치 기술세미나) 정형, 반정형, 비정형 등 데이터 형태, 데이터 활용 목적 등에 맞춘 사례 중심 기술 세미나를 지속 개최('18~)

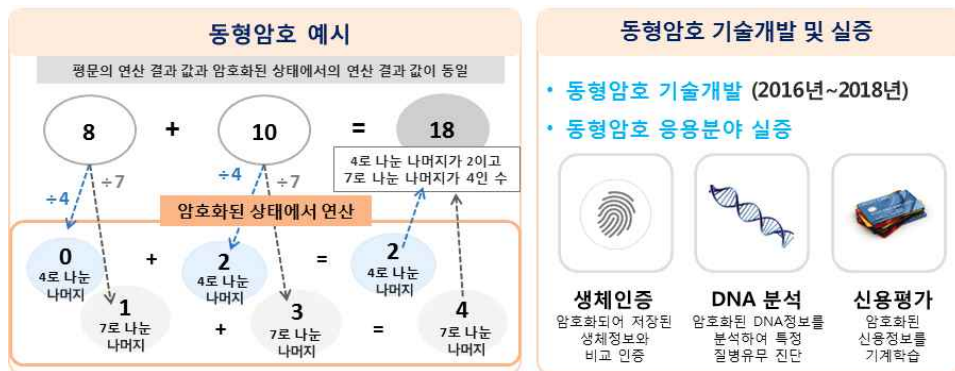
* 과기정통부·행안부·방통위·인터넷진흥원(KISA)·정보화진흥원(NIA) 공동 개최(3.29, 5.31)

- (데이터 보안 신기술 적용) 데이터의 위변조 방지를 위한 블록체인 기술, 암호화된 상태에서 데이터 활용이 가능한 동형암호기술 등 신기술 적용·실증 추진('18~'20)

* (블록체인) 개인정보 제공조건(대상·용도·가격), 접근권한 설정 등, (동형암호) 암호화되어 저장된 생체정보와 비교인증, 암호화된 DNA분석으로 질병진단, 암호화된 신용정보 평가 등

◆ 동형암호 기술현황

- [미국] IBM Helio(최초의 동형암호 기술)('12년 개발), MS SEAL('16년)
- [프랑스·스위스] 젤알토, 베르사유대 등 TFHE('17년)
- [한국] 서울대 해안(연산속도, 효율, 활용 분야 측면에서 기술 우위)('17년)



- 데이터 활용 스타트업, 중소·벤처기업의 블록체인 기술 적용 및 테스트를 지원하기 위한 블록체인 기술지원센터 설치·운영('19~)
- (데이터 주권 강화) 개인정보가 해외에서도 안전하게 보호되도록 국외 이전 중단 명령권, 국외 재이전시 보호조치 의무 신설 등 제도를 정비('18)
- EU GDPR('18.5)의 우리기업 피해 최소화를 위해 대응 가이드라인 마련 및 적정성 평가* 승인 추진('18)

* EU 제3국이 개인정보 보호를 위한 적절한 수준을 갖추고 있는지를 심사하여 제3국의 기업들이 별도 절차 없이 EU 시민의 개인정보를 이전·처리할 수 있도록 허용하는 제도
(적정성 평가 미 승인 시, 개별 기업 단위로 개인정보 역외이전 승인·보호조치가 필요하여 기업 비용 증가)

< 한-EU 적정성 평가 승인 로드맵 안 >

구분	1단계	2단계	3단계	4단계
내용	EU 집행위 초기결정	제29조 작업반 의견 제시	EU 각 회원국의 대표로 구성된 제31조 위원회 심의·승인	EU 집행위원회의에서 최종 결정
목표 시한	~'18.7월	'18.7~8월	~'18.9월	~'18.10월

[전략 2] 데이터 가치사슬 소주기 혁신

- ◇ 데이터 구축·개방(4차 산업혁명 핵심데이터 등, ~22) → 저장·유통(데이터 거래의 비즈니스화, '18~) → 분석·활용(산업·사회 혁신 활용, '18~) 등 소과정에 걸쳐 실제데이터 기반 영역별(의료·교통 등) 국가 빅데이터 지원체계 마련

1 양질의 데이터 구축·개방

- 4차 산업혁명의 핵심 기반인 산업별 실제데이터, AI 학습데이터를 조기구축하고, 공공·민간 데이터의 획기적 개방을 추진

① 4차 산업혁명 기반 실제데이터, AI 데이터 전방위 구축

- (산업별 빅데이터 구축) 산업적 활용도가 높은 주요 영역별로 기기·센서 등을 통해 수집·생성되는 현장 중심의 실제데이터 구축 확산('18~)

< 주요 분야 추진내용 >

분 야	빅데이터 구축 방향
의 료	의료 인공지능(AI) 개발을 위한 학습용 의료데이터셋 및 의료정보시스템 구축('18~)
제 조	스마트공장 내 스마트센서 기반 제조공정 빅데이터 구축('18~)
농·수산업	농식품 통합 데이터 지도 구축('18~), 어선정보 등 수산정보 제공('18~)
도 시	방범방재, 3차원 공간정보 등 도시운영 통합 플랫폼 기반 도시데이터 수집 기반 마련('18~)
교 통	자율차, 스마트도로 관련 차량·신호등 등 실시간 교통DB 구축('17~)
환 경	대기오염·수질오염 등 환경 빅데이터 통합·개방 플랫폼 구축('17~)

- 모든 공공데이터는 원시 데이터(raw data) 형태로 최대한 모으고, 이 중 민간 수요가 높은 데이터를 국가중점데이터로 선정('17. 48개 분야→22 128개 분야), 조기 개방 확대

◆ '18년 국가중점데이터 개방 분야(29개 예정)

- [혁신성장동력 육성] 인공지능 의료영상 소재정보, 자율주행 영상판독DB 등 6개
- [사회안전망 강화] 공공시설물 안전정보, 식품안전관리인증정보 등 7개
- [정부투명성 강화] 선거종합정보, 소비자안전정보 등 5개
- [일자리 창출] 해외시장 진출지원정보, 농업활동지원정보 등 6개
- [사회적 가치 확산] 민원분석정보, 도시재생 종합정보 등 5개

- 산업별 원시 데이터의 풍부한 수집·생성을 위해 빅데이터 전문센터*를 육성 (~'22, '18년 3개소)하고 각종 빅데이터센터간 협력 네트워크 확대('18, 35개→'22, 100개)

* 활용 가치가 높은 빅데이터를 수집·가공·개방하고 빅데이터 분석·활용을 전문 지원

◆ 미국 빅데이터 협력네트워크 사례

- o 미국 전역에 4개의 빅데이터 네트워크 허브를 구축하여 데이터 연계·활용 생태계 조성
 - 서부(의료), 중부(농업), 북동부(에너지), 남부(제조)
 - 250개 이상의 대학, 재단, 기업 등이 참여하여 데이터 수집·분석, 사회문제 해결 등 지원



- o (AI 학습용데이터 구축) 이미지·상식 등 범용 AI데이터셋을 우선 구축('17. 688만건*→'22. 1.1억건), 법률·특허·의료 등 전문분야 AI데이터셋을 수요 기반으로 단계적 구축·보급**('17. 211만건*→'22. 4,800만건)

* △상식 : 한국어 위키백과 정보 25만건, △이미지 : 한국인 안면이미지 648만건, 음식이미지 15만건
△법률 : 교통사고, 층간소음, 창업 인·허가 12만건, △특허 : 전기·전자분야 정보 199만건

** NIA(주관), ETRI, KIST, 특허정보원 등을 통해 추진(지능정보산업인프라조성사업, '17~'22 총500억원)

- 많은 수작업이 필요한 기계학습 데이터 前처리 과정(이미지 태깅 등)은 민간의 집단지성 플랫폼* 활용(수요 중소·벤처기업에 바우처 지원), 고품질 데이터 확보와 일자리 창출 연계('19~)

* (예) 국내 스타트업 크라우드웍스社가 운영 중인 AI 학습용데이터 구축 플랫폼에 약 3,000여명 참여 중

- 정부지원 AI R&D 및 시범사업 진행 과정에서 축적된 모든 AI 학습데이터를 'AI 허브('18.1)'를 통해 공개 추진('18~)

◆ 학습용 데이터 구축 방안

- o (수요조사) 매년 산업별로 필요한 기계학습용 데이터의 수요를 조사하여 직접 구축 ('22년까지 약 8.9백만건)
- o (R&D 축적) 인공지능 공공 R&D 과정에서 축적된 기계학습 데이터를 AI 허브에 집적하여 공개('22년까지 약 8.9천만건 개방)
- o (데이터 바우처 제공) 기업 수요에 맞춰 기계학습용 등으로 변환할 수 있도록 '데이터 바우처'를 3년간('19~'21) 총 600건 제공하여 가공비용 일부 지원

참고

국내 주요 데이터 현황

분야	기관명	주요 데이터	활용 사례
① 보건 의료	심사평가원/ 건강보험공단	▶ 수진자 인적 사항, 진료과목, 병명, 급여비용, 투약 정보 등	환자별 맞춤형 진단·치료 서비스 제공 및 정밀의료 솔루션 개발 등
	민간의료기관	▶ 환자 진료기록부, 의료영상 정보, 처방원무 정보 등	
② 교통	교통연구원/ 교통안전공단	▶ 도로현황, 도로시설물 관리, 교통량, 사고지점, 피해상황 등	교통 혼잡 완화를 위한 분석 서비스 및 교통사고 원인 분석 서비스 제공
	한국스마트카드	▶ 시간별 지역별 이동수단 승하차 인원 등	
③ 금융	한국은행/신용정보원	▶ 경제·금융 통계, 개인·법인 대출, 세금체납, 채무불이행, 부도 등	맞춤형 금융 서비스 개발 및 보험사기 분석 연체자 예측 모델 개발 등
	은행/보험/신용카드	▶ 계좌정보, 대출, 상품거래, 인터넷 뱅킹 이용 내역, 민원 가맹점 등	
④ 통신/ 미디어	통신사/ IPTV 업체	▶ 가입자/위치 정보, 유동인구, 서비스별 트래픽/구매 내역 등	감염병 차단 서비스, 상권 분석, 콘텐츠 추천 서비스, 광고 전략 개발 등
	언론진흥재단/ 방송광고진흥공사	▶ 종사자, 광고시장 현황, 구독/시청 패턴, 디지털 콘텐츠 매출 등	
⑤ 도시/ 공간	지자체(광역/기초)	▶ 도로/가로등 위치, 전력/가스 공급 체계, CCTV, 주차/횡단보도 정보 등	지능형 도시 서비스, CCTV 기반 보안 서비스, 입지분석, 부동산 가격 예측 등
	국토연구원/ 한국토지주택공사	▶ 지형, 산업입지, 택지, 부동산거래, 3차원 공간정보 등	
⑥ 에너지/ 환경	한전/지역가스/ 한국에너지공단	▶ 전력 판매, 전력 시설물, 전력 계량, 과금, 태양광 정보, 소비 패턴 등	에너지 공급 제어·관리 서비스, 전력소비 패턴 분석, 자연재해 예측 등
	기상청/수자원공사	▶ 기상정보, 도로날씨, 기후변화, 자연재해, 가뭄/지하수 정보 등	
⑦ 연구	과학기술정보연구원	▶ 바이오, 소재 등 기초과학 데이터 및 대형연구장비, 연구노트 등	신약 후보 물질 발굴, 신소재 연구 등에 활용
	참조표준데이터센터	▶ 건강지수, 수질, 인체치수 등	
⑧ 문화/ 관광	문화정보원	▶ 예술작품, 문화재, 역사자료, 문화 산업, 도서, 체육 정보 등	문화재 위험관리 분석, AI 기반 관광안내서비스 등
	한국관광공사	▶ 관광지, 다국어 관광정보, 숙박, 음식점, 축제 정보 등	
⑨ 제조	삼성전자, 엘지전자, 현대/기아자동차 등	▶ 자재관리·구매/생산/설비관리/품질관리 정보 등	제조공정 자동화, 재고관리, 신상품개발, 자원관리 등
⑩ 유통	백화점, 할인마트, 홈쇼핑 외식업체 등	▶ 고객현황, 구매이력, 배송이력, 식자재 내역 등	상품추천 등 고객마케팅, 반품률 예측 등

② Open Data : 공공·민간 데이터의 획기적 개방

- (공공부문 전면개방) 공공데이터의 원칙적 개방(국가안보·개인정보 제외) 등을 통해 공공부문 핵심데이터(통계·연구·참조표준 등) 구축·개방 전면화
 - 모든 공공기관이 보유한 공공데이터를 전수조사하고 데이터 현황을 보여주는 국가데이터맵을 구축하여 공공데이터 통합 관리체계 마련
 - * 현재 공공데이터 개방 2.5만개('18.6), 전수 조사 결과 공공데이터 중·장기 개방 계획 수립
 - 국가안보·개인정보를 제외한 전력·환경·안전 등 국가중점데이터 개방 확대
 - 통계, 연구, 참조표준 등 분야별 공공데이터의 체계적 구축·개방 추진

분 야	주요 추진방향
통계 데이터	▲ 통계빅데이터센터 확대('18, 3개소), ▲ 원자료에서 오류 제거한 마이크로데이터 확대 * 마이크로데이터 : 통계조사 원자료에서 오류를 제거하여 작성한 기초데이터('17. 289종 개방, 매년 15종 이상 개방)
연구 데이터	▲ 바이오·소재·연구장비·AI 우선구축('18~), ▲ 관리계획 제도화 등 제도정비('18~)
참조표준 데이터	▲ 연구산업현장 측정데이터 참조표준 확대 ▲ 수질·건강지수 등 분야별 센터 확대 * ('18.3) 재료, 물리, 화공 등 10개 분야 4.3만개 / 출연연, 대학 등 27개 센터 운영 중

- 기계학습이 용이하도록 개방 공공데이터의 오픈포맷(3단계) 비중을 지속 확대하고, 데이터의 손쉬운 가공을 위해 개방표준 지정 확대(~'22), 공공데이터 품질관리 고도화* 등을 통해 소공공기관의 데이터 품질 대폭 개선

○ (오픈포맷 비율) ⁽¹⁷⁾ 78.8%→⁽¹⁸⁾ 85%, (개방표준서식) ⁽¹⁷⁾ 109종→⁽²²⁾ 200종

○ (품질관리 수준평가제) ⁽¹⁸⁾ 중앙행정기관→⁽¹⁹⁾ 지자체→^(20~) 공공기관

* 공공데이터 관리지침(행안부 고시 제2016-42호, 2016.11)을 제정하여 공공데이터 품질 관리 일환으로 '공공데이터 수준 평가제도' 기준 및 원칙 제시

- (공공재 성격 민간데이터 개방) 유통·금융 등 민간의 자율적 데이터 개방*을 확대하고, 공공성이 높은 민간데이터(예: 코스콤 주식시세)에 대해서는 연구·창업 목적 시 부담없이 활용할 수 있도록 중재·개방 지원**('18.下~)

* 산업적 활용도가 높은 민간 빅데이터 개방 : SKT(통신), 네이버(검색), 언론진흥재단(뉴스), KB카드(카드매출), CJ올리브네트웍스(유통), 신한은행(금융거래), 아산병원(의료영상) 등

** 구매바우처 지원, 공익데이터 연구반 운영 → 가격인하, 선활용-후정산 등 권고

- 데이터의 효율적인 저장·관리를 위한 클라우드 본격 확산 및 데이터 유통 촉진을 위해 민간·공공을 연계한 개방형 데이터 기반 구축

① 클라우드 기반 데이터 관리 확대 및 안전성 제고

- (규제완화) 민간 클라우드서비스 이용기관을 공공기관에 한정하지 않고 지자체 등으로 확대(부처협의로 관련 제도 개선 추진)

* 공공분야 클라우드 이용 활성화를 위해 해커톤 개최(4.3~4) 등 사회적 합의 도출 지속 추진

- (All@Cloud 확산) 스마트시티, 스마트공장, 창업 등 국민 체감도가 높은 분야에 수요 맞춤형 지원으로 클라우드 서비스 접목·확산

* (예) 스마트시티@Cloud, 스마트공장@Cloud, 창업@Cloud 등

- '22년까지 중소·벤처기업 1만개 이상에 클라우드 도입을 지원

* (이용료 지원범위) 산업단지 내 기업 → 전 사업체

- (보안강화) 공공기관이 이용하는 민간 클라우드의 안전성·신뢰성을 강화하기 위해 클라우드 보안인증제 확대(IaaS → SaaS 추가, '18)

- (선도활용) 클라우드 이용효과가 큰 분야*를 발굴하여 클라우드로 전환을 추진하고, 금융·의료·교육 등 주요 서비스 산업의 클라우드 선제적 도입을 촉진하기 위한 클라우드 선도활용 시범지구를 조성('18~)

* (사례) 국가R&D, 초중고 SW교육, 공공기관 스마트협업, 대학 클라우드, 스마트 문화재 서비스, 스마트시티 통합플랫폼, 클라우드 공인인증서 등

- (개방형 플랫폼 확산) 국내 클라우드 플랫폼인 '파스-타(PaaS-TA*)'의 기능을 고도화하고 오픈소스를 중소·벤처기업에 제공

* 사실상 세계 표준인 Cloud Foundry의 오픈소스를 기반으로 국내에서 개발한 클라우드 플랫폼으로, 코스콤·SK C&C·KT 등 10여개 기업·기관에서 사용 중

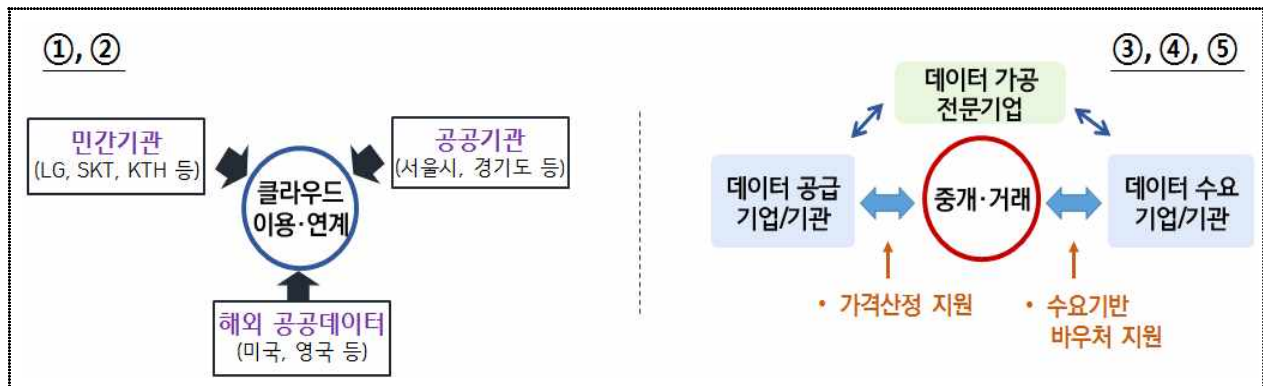
② 개방형 데이터 거래 체계 구축 등을 통한 양질의 데이터 유통 촉진

○ (데이터 거래 기반 구축) 누구나 데이터를 한 곳에서 쉽고 빠르게 등록·검색·거래할 수 있도록 민간·공공을 연계한 데이터 거래 기반 구축

- ① 민간·공공 데이터 포털을 연계, 개방형(CKAN*)으로 고도화
- ② 국내·외 주요 데이터 플랫폼 연계
- ③ 데이터 가공 전문기업과 수요기업 매칭
- ④ 초기시장 활성화를 위해 중소·스타트업에 데이터 바우처(구매·가공비용 등 지원) 제공
- ⑤ 바우처를 통해 가공된 데이터는 상품으로 다시 등록·판매

* Comprehensive Knowledge Archive Network : 오픈소스 기반 플랫폼, 美·英 등 31개 중앙정부 등 사용

< 데이터 거래 기반 구축 추진방안 >



○ (민간주도 데이터 거래 촉진) 데이터상품(날씨, 지도, 기업정보 등) 보유기업들이 협력하여 데이터를 공유 판매하는 데이터 거래소를 설립·운영시 초기단계 지원(18~)

* 공공데이터 가공, 오픈소스 분석SW 자원, 비식별처리 솔루션 등 연계 지원

(KTH, CJ올리브네트웍스 : api 거래 플랫폼 운영 중 / 빅데이터센터 협약체 : 통합 빅데이터 거래소 구축 검토)

○ (제도정비) 국민들이 데이터를 합리적인 가치로 자유롭게 거래할 수 있고, 데이터 독점 등 불공정행위를 방지할 수 있도록 제도 연구·정비(18~)

공정경쟁 환경조성	지재권 보호·활용 촉진	데이터 거래 투명성 확보
·데이터 기업 간 M&A 시 야기 되는 경쟁제한 효과에 대해 심사기준 연구(공정위, '18)	·AI 창작물 보호, 빅데이터 수집 가공 시 저작권 침해 등 지재권 이슈 논의·검토(문체부특허청 '18~)	·데이터 거래 가이드라인 및 표준계약서 개발·보급 (과기정통부, '18~)

참고

해외 주요국 데이터 거래소 사례

□ [중국] 세계 최초로 민관 합작형태의 데이터 거래소 설립을 통해 공공데이터의 가공·판매, 민간 데이터 거래 등 공공 주도 형태

- 중국 정부의 빅데이터 산업발전 클러스터 구축계획에 의거, 공업신식화부의 승인을 받아 구이양 빅데이터 거래소 설립('15.2)
 - 거래소는 총 자본금 5,000만 위안(약 84억원) 중 정부가 국유자본 36%를 투자하고, 5대 기업들이 각각 출자한 민관 합작 법인으로, 총 100여명 근무
- 자체 유통 플랫폼인 GBDE(trade.gbde.com)를 운영중이며, 현재 등록 데이터는 3,643건으로, 공공데이터가 80%를 차지
 - * 공공데이터를 무료로 개방하는 他국가들과 달리 중국은 정부 소유의 공공데이터 가공·판매
- 철저한 회원제로, 현재 공공·민간 300여개 회원사를 보유중이며, 거래소 설립 이래 누적 총 거래액은 '18.5월 기준 3억 위안(약 500억원) 수준

□ [미국] 개인정보 활용이 비교적 용이한 환경하에서 세계 최대 규모의 데이터 브로커 시장(약1,500억불) 등 민간 기업 위주로 데이터 거래 활발

- ①세계 최대 데이터 브로커기업*이 데이터 가공분석을 통해 맞춤형 서비스 제공이 활성화 되어 있으며, ②자체 플랫폼을 통해 국가별 통계 데이터를 시각화로 가공, 전 세계에 판매하는 기업도 성업 중**
 - * 엑시움社は 데이터 가공분석을 통해 금융·유통회사 등에 적합한 맞춤형 서비스 제공(연 8억불)
 - ** 클릭社は 자체 플랫폼(QlikSense, QlikView)을 통해 주요 35개국 주식시세정보, 기업재무정보, 환율정보, 인구·산업통계정보 등을 시각화로 가공하여 전 세계로 판매

□ [일본] 정부차원에서 종합계획을 설립하여 주도하는 것이 특징으로, '20년까지 IoT로 수집된 데이터 거래를 위한 유통 시스템 구축 준비 중

- 일본은 '16년을 4차 산업혁명 원년으로 삼고, 데이터 활용촉진을 위한 환경정비 등 7대 추진전략을 마련 등 정부 차원에서 종합계획을 설립·추진중
 - 특히, 히타치, NTT 등 100개 기업이 참여하고 총무성이 지원하는 『IoT 데이터 거래소』 개설 추진('20년)

- 데이터를 활용한 산업적·사회적 가치 본격 창출을 통해 산업 전반에 성장활력을 제고하고 국민 삶의 질 향상에 기여

① 데이터 기반으로 산업경쟁력을 제고

- (중소기업 혁신) 성장한계에 직면한 전통 중소기업과 빅데이터 전문기업을 매칭, 新제품 개발·맞춤형 홍보 등 혁신 돌파구 마련 지원(500개사, ~'22)

* (예) 티젠(茶 생산업체) : 북미 소비자 맞춤형 차 출시 → 해외시장 진출 한계 극복(30만불 수출, '17)
HY스타일(발관리제품생산) : 타겟고객층 맞춤형 홍보 → 매출 증대(전년대비 19% ↑, '17)

- (연구데이터 사업화) 바이오, 미래소재, 대형연구장비 등 연구데이터 집약형 분야에서 산업적 활용을 촉진하기 위한 선도 프로젝트* 실시('18~)

* △ 바이오: 화합물 데이터(약 50만건) 활용 신약 개발 플랫폼 구축(시간·비용 1/3 ↓),
△ 미래소재: 열전소재 물성 빅데이터 플랫폼 기반 소재 탐색 최적화 기술 개발,
△ 대형연구장비: 대규모 연구데이터 공유·교차 분석 기반 단백질 구조분석 연구 추진

② 빅데이터 활용 사회문제 해결 강화

- 데이터 기반의 국가적 중요 의사결정(다부처 정책 수립 및 사회현안 해결) 지원을 위한 '공공빅데이터센터' 설치

- 빅데이터 기반의 실시간 모니터링, 사전예측, 조기대응 등을 활용, 사회문제 해결을 위한 빅데이터 플래그십 프로젝트 투자를 대폭 확대*

* 빅데이터 플래그십 프로젝트 투자확대 목표 : ('18) 8건 → ('19) 20건 → ('22) 25건

분 야	주요 추진내용
건강 저해요인 차단	미세먼지, 감염병 → IoT 기반 빅데이터 수집 및 실시간 모니터링·예보·원인규명 등 선제 대응체계 마련(과기정통부·환경부·복지부, '18~)
재난안전 대응 강화	긴급구조, 노후건물 붕괴, 전기화재 → 정밀측위, 지능형CCTV 데이터 수집·분석 등 대형재난 신속대응 및 예방 강화(과기정통부·행안부·국토부, '18~)
디지털 성범죄 피해 방지	지진, 화산 → 지진파 데이터 구축, 화산특화연구센터 개소(기상청, '18~)
빅데이터 분석을 통해 몰래 카메라 등 디지털 성범죄 우범지역 도출·제공을 통해 사고 및 피해 방지	(과기정통부·경찰청, '18~)
복지 사각지대 해소	고독사, 자살 → 소득·채납정보 등 빅데이터 기반 사회취약계층 발굴 및 맞춤형 사회보장 서비스 제공(복지부, ~'22) 장애인 편의·안전 → 특별교통수단 운영기관(지자체)이 이용패턴 분석을 통해 효율적 이용 가능토록 활성화 유도(국토부, '18~)
고품질 의료서비스 제공	진료정보 이동 불편 → 의료기관 간 진단·투약·영상정보 등 온라인 진료정보 전자교류 서비스 전국 확대(복지부·과기정통부, ~'22)

[전략 3] 글로벌 데이터산업 육성기반 조성

- ◇ 빅데이터 산업이 4차 산업혁명을 선도하는 핵심 성장주체로 거듭날 수 있도록 4차 산업혁명 요소기술 융합, 분석 전문인력 양성, 컴퓨팅파워기반 기업성장 인프라 지원 등을 아우르는 역동적 산업 생태계 조성

1 빅데이터 관련 선도기술 조기 확보 : 선진국 대비 90% 이상 수준

- (핵심기술 확보) 초연결지능화, 분석 신뢰성 향상, 활용도 제고 등 빅데이터 기술발전 전망을 고려하여 핵심 원천기술 개발 추진('18~'22)
- 빅데이터 관련 △분석(오픈 데이터 플랫폼, 엣지 분석 기술, 공간 빅데이터 분석 기술 등), △인공지능(상황인지형 데이터 관리, 초대규모 모사현실 기술 등), △초연결지능화(초연결 데이터 관리, 현실-가상공간 실시간 연계 기술, 지능형 집단협업 플랫폼 기술 등) 분야 R&D 지원 강화

< 빅데이터 기술개발 로드맵 (안) >

1단계 ('17-'18)	2단계 ('19-'20)	3단계 ('21-'22)
분석 중심	인공지능 중심	초연결지능화 중심
<ul style="list-style-type: none"> · 오픈 데이터 플랫폼 · 엣지 분석 기술 · 개인정보 안전한 처리 · 데이터 증식 	<ul style="list-style-type: none"> · 상황인지형 데이터관리 · 초대규모 모사현실 · 실시간 지능형 분석추천 · 빅데이터셋 이상감지 	<ul style="list-style-type: none"> · 초연결 데이터 관리 · 지능형 집단협업 플랫폼 · 직관적 의사결정 분석

- (관련기술과 융합) 초대용량 데이터의 고속 수집·처리·분석·예측기술 및 빅데이터 기술과 4차 산업혁명 요소기술(AI·IoT·클라우드·정보보호)을 연계한 융합 선도기술 개발 가속화('18~'22)

< 빅데이터 및 관련기술 융합 추진 >



- (청년일자리 연계) 대학 졸업예정자 등 청년 대상 실무 중심의 빅데이터 전문 프로그램*을 지원하고 교육이수자를 양질의 일자리에 연계(~'22, 9천명)
 - * 빅데이터 청년인재 일자리연계 사업(과기정통부), 4차 산업혁명 선도인력 양성사업(고용부) 등
- (고급인재) 데이터 분석 전공 운영지원(SW중심대학, 교수·시간강사 연수 등), 빅데이터 전문 연구센터(ITRC·ERC) 확대('17, 1개→'22, 6개) 등을 통해 데이터 과학자 집중 양성(~'22, 8천명)
- (실무인력) 산업 수요기반 전문인력 양성을 위해 재직자 대상 실무교육 지원*, 국가공인 데이터 자격제도**를 활성화하고 데이터 분석 국가기술자격제도(빅데이터 분석기사)를 신설('19)·운영(~'22, 33천명)
 - * 데이터 아카데미(과기정통부), 국가인적자원개발컨소시엄(고용부) 등
 - ** 설계(데이터 아키텍처)·개발(SQL)·분석(데이터 분석) 분야 '17년까지 12천명 배출
- (데이터 경연장) 각 분야 빅데이터센터가 데이터셋과 문제를 제시하고 학계, 스타트업, 연구기관 등이 참여하여 데이터 연계 분석을 통해 새로운 가치를 찾아내는 데이터 인재발굴 플랫폼* 운영('19~)
 - 데이터 기부 및 문제해결 상금지원, 비식별화 지원 기능 등을 구현하고 데이터 안심존과 연계하여 데이터분석·알고리즘개발 편의 제공
 - * 캐글(Kaggle) 방식으로 검토·추진(캐글 : '10년에 설립된 전세계 데이터 과학자가 참여하는 집단지성 기반 문제해결 플랫폼으로 '17. 3월에 구글에 인수 / 기업 및 단체에서 데이터와 해결과제를 등록하면 데이터 과학자들이 이를 해결하는 모델을 개발하고 경쟁)

< 데이터 인재발굴 플랫폼 안 >



- (인프라 지원 강화) 빅데이터 개발에 필수적인 컴퓨팅 자원, 데이터셋 등이 집적된 판교 글로벌 ICT 혁신 클러스터 지속 고도화 및 활용 지원
 - * K-ICT빅데이터센터, K-ICT클라우드혁신센터, 글로벌 IoT시험인증센터, 정보보호클러스터, HPC이노베이션허브 등 / 세계 최고수준의 성능을 갖춘 초고성능컴퓨터 5호기와도 연계

< 판교 글로벌 ICT 혁신 클러스터 추진방향 안 >



데이터 스타트업 육성 HUB

- (맞춤형 성장 지원) 데이터 분야 스타트업·강소기업을 중점 발굴·육성 (~'22, 100개사)하고 유망한 데이터기업의 해외진출을 밀착 지원(~'22, 80개사)
 - K-Global DB-Stars* 사업을 통해 데이터 활용 사업화 아이디어를 보유한 스타트업을 지속 선발하고 컨설팅·투자유치 등 지원 확대
 - * '14년부터 지원된 데이터 분야 정부 대표사업으로, 그간 55개 스타트업 발굴하여 국내·외 340억원 민간투자유치, M&A 2건, 국제수상·해외진출 27건 등 성과 창출
 - 공공데이터 활용 창업 콜라보(Collabo) 프로젝트*를 통해 민·관이 협업, 맞춤형 창업 지원으로 성장 가능성이 있는 데이터 강소기업 육성
 - * 아이디어 발굴부터 자금지원, 해외진출 등 공공데이터를 활용한 창업 쏠단계에 걸친 종합지원 프로그램(행안부·중기부·특허청·기본·신보·엔젤투자협회 등 10개 기관 참여)
 - 국내 데이터 기업의 솔루션에 대해 진출 가능성이 높은 국가를 선정하여 수요기업의 솔루션 현지화·마케팅, 해외전시회 참가 등 지원
 - * ('19) 동남·중앙아시아 → ('20) 중동 → ('21) 동유럽 → ('22) 중남미 → ('23) 아프리카 지역

V. 기대효과



□ 산업·기업의 혁신성장 가속화

- 데이터경제의 활성화는 조선, 자동차, 철강 등 기존 주력산업의 생산성과 효율성을 향상시켜 재도약의 발판을 제공하고,
- 금융, 유통·물류, 의료·보건, 바이오 등의 산업을 새로운 성장동력으로 자리매김과 함께 스타트업과 벤처기업의 질적 성장을 도모

□ 국민 개개인의 행복 증진

- 데이터 활용에 기반한 생활밀착형 서비스는 일상생활에서 발생하는 사회문제를 해결하고, 삶의 질과 연관된 건강, 안전, 편의 등을 증진
- 국가에 대한 의무, 공공복지 등에서 국민 누구나 소외되지 않고 공정한 기회를 가질 수 있는 사회 실현

□ 국민이 신뢰하고 공감하는 정부

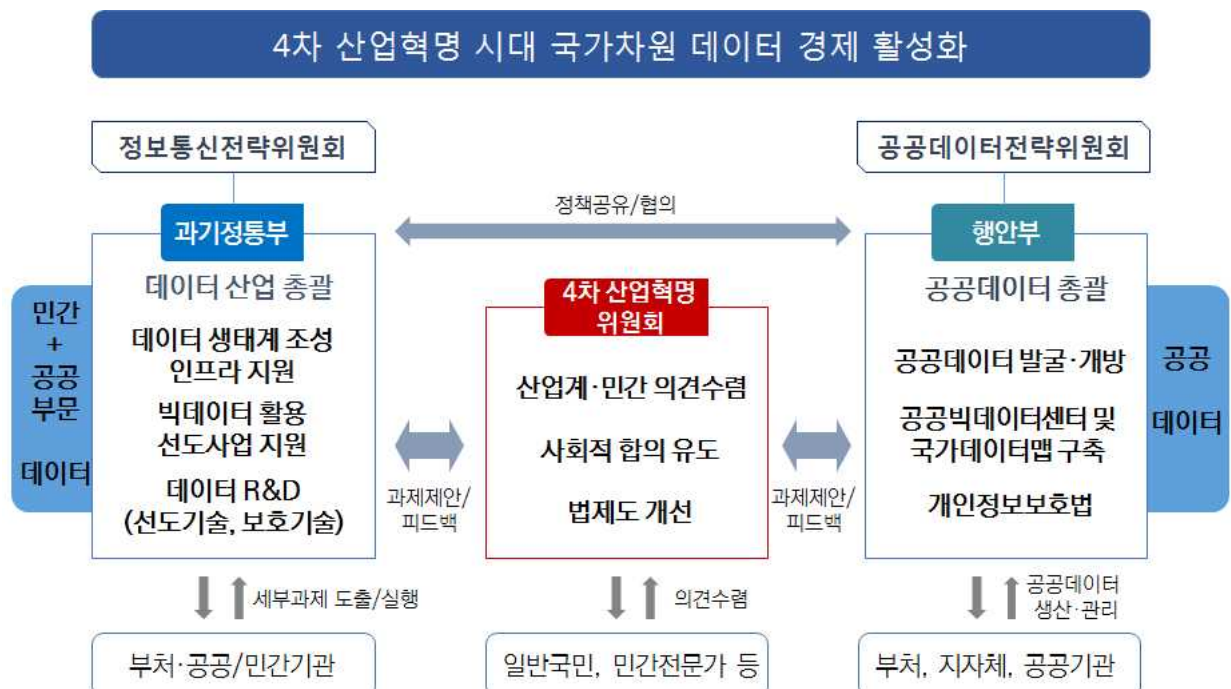
- 국가기관, 행정/공공기관의 데이터 기반 의사결정과 데이터 기반 디지털 행정서비스는 정책에 대한 일관성과 신뢰성을 확보하고 국민의 체감도를 제고

⇒ 4차 산업혁명 시대에 전국민이 데이터경제를 통한 경제적 과실을 체감하고 함께 나누는, 데이터 기반의 혁신성장 선도국가로 자리매김

VI. 추진체계

- 데이터 기반으로 4차 산업혁명 선도국가 반열에 오르기 위해 국가차원 데이터 경제 활성화를 위한 범국가적 대응체계 구축 추진
 - (4차산업혁명위) 국가적 데이터 경제 정책방향의 산업계·민간 의견수렴 및 사회적 합의 유도(정보통신전략위·공공데이터전략위 협업)
 - (과기정통부) 데이터 경제 활성화 대응 관련 전략 수립 및 핵심과제 도출, 정책 실행 등의 데이터 산업 총괄 역할 수행
 - (행정안전부) 국가사회 전반에서 공공데이터 수집·개방·활용 추진 전략 수립 및 핵심과제 도출, 공공 빅데이터 활용 활성화 정책 실행 등의 공공부문 총괄 역할 수행
 - (관계부처·공공기관) 소관 산업 분야별 빅데이터 활용 정책과제 발굴 및 시행, 빅데이터 구축 확산

< 데이터 경제 활성화 추진체계 (안) >



< 과제별 추진 로드맵 >

번호	추진 과제 및 일정	추진 부처
	2018년 2019년 2020년 2021년 2022년	

[전략1] 데이터 이용제도 패러다임 전환

①	데이터 이동권 확립 : 국민 데이터 주권 찾기(MyData) 금융·통신 에너지 의료 유통	과기정통부, 금융위 등
②	개인정보 안전한 활용 촉진 법개정 데이터 안심존 비식별조치 콘테스트·기술세미나 블록체인·동형암호 적용·실증 GDPR승인	행안부, 방통위, 과기정통부

[전략2] 데이터 가치사슬 쏠주기 혁신

①-1	4차 산업혁명 기반 실제데이터, AI 데이터 전방위 구축 산업별 빅데이터 구축 빅데이터전문센터 육성 원시 데이터 축적 국가중점데이터 선정·개발 AI 학습용 데이터 구축·개발 AI 학습용 데이터 전처리 민간플랫폼 활용	과기정통부, 행안부 (민관 관련기관)
①-2	Open Data : 공공·민간 데이터의 획기적 개방 공공부문 4대 핵심데이터 구축·개발 개방표준 확대 및 데이터 품질고도화 공공재 성격 민간데이터 개방	행안부, 과기정통부, 통계청, 산업부
②-1	클라우드 기반 데이터 관리 확대 및 안전성 제고 보안인증제 확대/법개정 All@Cloud 확산 중소·벤처기업 클라우드 도입 지원 클라우드 선도활용 시범지구 조성	과기정통부, 행안부
②-2	데이터 거래 기반 구축 등을 통한 양질의 데이터 유통 촉진 데이터 거래 기반 구축 민간 데이터 거래소 초기단계 지원 공정경쟁 환경조성 지재산 보호·활용 데이터 거래 투명성 확보	과기정통부, 공정위, 문체부, 행안부, 특허청
③-1	데이터 기반으로 산업경쟁력을 제고 중소기업·빅데이터 전문기업 매칭 신약개발 미래소재 대형연구장비	과기정통부, 산업부
③-2	빅데이터 활용 사회문제 해결 강화 미세먼지·감염병 재난안전 디지털 성범죄 복지 사각지대 고품질 의료	과기정통부, 환경부, 복지부, 국토부, 행안부

[전략3] 글로벌 데이터산업 육성기반 조성

①	빅데이터 관련 선도기술 조기 확보 분석 기술 인공지능 연계 기술 초연결지능화 연계 기술 융합 선도기술 (IoT Cloud AI Security)	과기정통부, 국토부 등
②	미래수요 대응 전문인력 확충 빅데이터 전문프로그램 대학연구센터 재직자 실무교육 데이터 자격제도 데이터 인재발굴 플랫폼	과기정통부, 고용부, 통계청
③	빅데이터 전문기업 성장 지원 판교 글로벌 ICT 혁신 클러스터 고도화·활용지원 맞춤형 성장지원(K-Global DB Stars 등) 해외진출 밀착지원(K-Global DB Stars 등)	과기정통부, 행안부

◆ 산업별 빅데이터 구축·확산·Best Practice 도출 등은 [붙임] 분야 중심으로 소관부처에서 추진, 필요시 4차위에 안건으로 보고

① 금융

금융 산업의 경쟁력 제고를 위해 신용정보, 예금·대출 등 은행업 핵심 데이터 수집 및 개방·유통 확대 지원

- 신용정보원, 보험개발원에 집중된 신용정보를 활용한 표본 DB* 구축을 통해 금융분야 빅데이터 테스트베드 우선 추진(금융위)
 - * 공공부문의 신용정보 인프라로서 대형사와 창업·핀테크 기업간 정보격차 해소 기대
- 클라우드 기반 개방형 플랫폼을 통해 은행업 핵심 데이터를 구축·개방(18~)
 - * 비식별 조치 데이터(예금·대출 금액, 급여·세금·보험료 이체, 고객부가정보 등), 기업 데이터(기업명, 매출액, 연체·부도여부 등), 통계 데이터(성·연령·거주지별 금융·부동산 자산, 저축액, 부채현황, 월소비액 등)

◆ 주요센터 : 신한은행, KB국민카드, 삼성카드, IBK기업은행 등

② 통신·미디어

산업적 활용도가 높은 통신사, 인터넷포털사 등이 보유한 민간 데이터(위치데이터, 검색데이터 등)에 대해 개방·유통 활성화

- 빅데이터 허브를 통해 데이터 공개 및 이중 데이터 결합지원, 융합·분석 서비스 제공(SKT, '17년말 기준, 데이터 979건)
- 데이터랩을 통해 분야별 인기검색어, 검색어 트렌드, 지역별 관심도 등의 데이터와 함께 지역·업종별 카드사용통계* 제공(네이버)
 - * 카드사(BC카드) 데이터와 네이버 검색어 데이터 융합·개방 첫 사례
- 각종 뉴스/미디어 정보를 산업에 활용 할 수 있도록 빅카인즈*를 통해 개방하고 소셜·방송·공공데이터 등 연계활용 확대(언론진흥재단)
 - * 42개 매체 약 4천만건 뉴스/콘텐츠를 빅데이터화 하여 관계망·워드클라우드·이슈트렌드 분석, 데이터 다운로드 및 시각화, 공공데이터 연계기능 등 제공

◆ 주요센터 : 빅데이터허브(SKT), 데이터랩(네이버), 한국언론진흥재단 (bigkinds.or.kr) 등

③ 보건 의료	민감한 의료정보의 특성을 고려한 사회적 대화 추진, 의료기관별 원내 의료데이터 활용기반 마련 및 수요자 맞춤형 익명 데이터 사업 확대
	<ul style="list-style-type: none"> ○ 시민사회, 전문가 등 사회 각계각층이 참여하는 보건의료 빅데이터 거버넌스 구성 및 입법 논의를 통해 데이터 보호·활용 원칙 수립('18~, 복지부) ○ 인공지능 학습용 고품질 데이터셋 구축·개방 및 임상 현장 시범 적용 등 의료 인공지능 테스트베드 조성('18~) <ul style="list-style-type: none"> * 정형, 비정형 의료데이터의 비식별화, 대용량 의료데이터클리닝, 경제적 데이터 라벨링, 오픈데이터베드 플랫폼 구축·운영 등 ○ 한의학 분야 환자 진료·질병 정보인 전자의무기록 데이터를 축적하고 이와 연계한 개인건강기록(PHR) 플랫폼을 구축·활용(한의학연) <ul style="list-style-type: none"> - 의약품 안전정보*를 종합 분석·관리하는 의약품 통합정보시스템 구축(식약처) <ul style="list-style-type: none"> * 임상시험, 허가·심사, 지도·점검, 이상사례 등 ◆ 주요센터 : 헬스이노베이션 빅데이터센터(서울아산병원), 암빅데이터센터(국립암센터), 보건의료빅데이터개방시스템(opendata.hira.or.kr) 등
④ 농수산	과학영농 확충, 해양수산 선진화를 위해 빅데이터 기술 활용 기반 구축 및 수요자 맞춤형 빅데이터 플랫폼 구축
	<ul style="list-style-type: none"> ○ 농식품분야 보유 데이터를 총 망라하여 데이터 지도*를 구축하고 빅데이터 융·복합 모델** 제공(농림부) <ul style="list-style-type: none"> * 농식품분야에서 보유 중인 데이터 현황을 전수 조사해 민간 활용도가 높은 데이터를 중점 수집·관리하고 활용분야별 데이터 연관관계를 시각화하여 제공 ** 귀농귀촌 맞춤지원, 농산물 유통센터 최적지 선정, 과수 작목전환 도우미 등 ○ 농식품 빅데이터 생태계 조성을 위한 핵심데이터 수집·제공 및 분석·활용을 위한 빅데이터 플랫폼 구축(농림부) <ul style="list-style-type: none"> * 농경지 전자지도(팜맵) 기반에 공공·민간정보(카드·통신사), IoT정보(스마트팜) 융합 ○ 주요 어선·업종별 경험정보, 현장관측 자료 등을 통합하여 새로운 수산자원 모델 등 수요자 맞춤형 수산정보 제공(국립수산과학원) ◆ 주요센터 : 농수산물유통공사, 국립수산과학원 등

⑤ 유통·물류	대·중소 유통업체가 공동 활용 가능한 상품 및 구매정보, 물류 정보 등에 대한 빅데이터 구축·활용 확대
<ul style="list-style-type: none"> ○ 정부 플래그십 사업 등을 통하여 식자재, 식품 등 유통 데이터를 수집·판매하는 데이터 거래플랫폼* 육성 지원 확대 <ul style="list-style-type: none"> * CJ올리브네트웍스는 CJ프레시웨이, 대한통운 등 그룹사가 수집관리하는 2만개 식자재, 전국 택배 유통정보 및 수요 정보(CJ멤버스십, 유동인구, 메뉴 등), 소셜 데이터 등 개방 및 거래('18.3 예정) ○ 국가 물류정책을 효율적으로 수립하고 기업의 물류 시스템 운영지원을 위해 전국 화물 빅데이터 DB* 구축 및 제공** 확대('18~, 국토부) <ul style="list-style-type: none"> * 화물차 통행실태, 물류거점별 진출입교통량, 출하실적, 물류시설, 운송수단 등 ** 국가교통 DB 플랫폼, 통계청 사업체 공간 DB 등 플랫폼을 통하여 데이터 제공 ○ 우체국이 보유한 물류배달, 쇼핑 데이터 등을 수집·분석하고 창업기업에게 개방하는 우체국 빅데이터 센터를 구축('18~, 우정사업본부) <ul style="list-style-type: none"> ◆ 주요센터 : CJ올리브네트웍스(bigdata.cj.net), GDS컨설팅, 우정사업본부 등 	

⑥ 제조	제조 현장의 빅데이터 활용성 제고 및 기업의 경쟁력 강화를 위한 제조 빅데이터 플랫폼 구축·활용 확대
<ul style="list-style-type: none"> ○ 민·관협력을 통해 스마트공장 구축을 2만개로 확대(~'22)하여 제조 공정 데이터 DB 구축·활용 확대(중기부) <ul style="list-style-type: none"> * 클라우드 기반의 스마트공장 구축을 추진하고 클라우드에 수집된 빅데이터를 활용하여 기업의 성과와 직결된 데이터(불량원인분석, 공정효율화 등) 제공 ○ 개별공장 차원의 데이터 활용이 아닌, 중소 제조업의 데이터 수집·공공 활용하기 위한 중소 스마트제조 데이터 센터 구축('20, 중기부) <ul style="list-style-type: none"> - 타이어사업에서 생산공정 및 설비용 각종 센서데이터와 운전데이터를 수집·분석하고 시각화하는 빅데이터 기반 스마트공장 구축(효성) ○ 제조현장에서 발생한 데이터를 축적·활용할 수 있는 제조 빅데이터 센터를 구축하여 품질불량 사전예측, 에너지 최적화 등 실현(LS그룹) <ul style="list-style-type: none"> ◆ 주요센터 : 삼성전자, LS그룹, 효성, 전자부품연구원 등 	

⑦ 도시	도시운영 효율화 및 문제해결 지원을 위한 IoT 기반 공공시설물, 전력 등 센싱 데이터 실시간 수집·분석 체계 구축
<ul style="list-style-type: none"> ○ 방법·방재 등 단절된 개별 도시정보시스템을 상호 연계한 도시운영 통합 플랫폼을 확산(~'22, 80곳)하여 도시데이터 수집 기반 마련(국토부) ○ 주요 도시의 고정밀 3D 공간정보 데이터를 구축·개방하고 위성정보 활용센터 설립하여 위성데이터 수집·이용 확대('19, 국토부) ○ 대용량 공간정보 등*을 취합·분석할 수 있는 공간빅데이터 플랫폼을 구축하여 소방안전 지도 구축 등 합리적 정책결정 지원('18~, 국토부) <ul style="list-style-type: none"> * 공간정보(지형도, 지적도, 위성영상 등), 행정정보(건축물대장, 지역별 인구현황, 버스노선 정보 등), 민간정보(유동인구자료, 신용카드매출, 교통카드 이용정보 등) ○ 에너지 컨설팅, IoT를 활용한 에너지 관리시스템 등 다양한 신사업 활성화를 위해 전력빅데이터센터의 데이터를 확대 개방('18~, 한전) <ul style="list-style-type: none"> * (기존) 6개월분 전력사용량 → (개선) 2년분 전력사용량·전기요금제·요금결제내역 ○ 데이터 기반 스마트시티 통합운영 플랫폼 연구개발을 통해 다양한 스마트시티 솔루션 모델 제공('18~) <ul style="list-style-type: none"> * 국가전략 스마트시티 R&D 사업으로 추진(국토부·과기정통부, '18~'22, 총 1,159억원) <p>◆ 주요센터 : 공간빅데이터(kobigs.molit.go.kr)(국토부), 전력빅데이터센터(한전), 국가공간정보포털(ksdi.go.kr)(국토부), V월드(map.world.kr)(공간정보산업진흥원) 등</p>	

⑧ 교통	교통카드, 네비게이션 정보 등 교통 빅데이터 통합·연계 플랫폼 구축·확산을 통한 실시간 교통대응 지원
<ul style="list-style-type: none"> ○ 공공·민간 교통 데이터*를 수집하여 차량, 대중교통 등 실시간 통행량 분석에 활용하는 교통 빅데이터 시스템** 구축(교통연) <ul style="list-style-type: none"> * (공공) 디지털 운행 기록계, 교통카드, 교통량, (민간) 통신, 네비게이션 정보 등 ** 전국 2차선 이상 도로의 교통량, 속도를 수집 분석하는 빅데이터 플랫폼을 고도화 ○ 전국 모든 도로에 센서를 부착하여 교통 혼잡정보 등을 IoT 연계 플랫폼 기반 실시간 수집하는 시스템을 구축하고 교통사고 원인분석 등에 활용 ○ 자율주행차 데이터의 대규모 축적·활용을 위한 데이터 공유센터를 구축('18~, 국토부) <p>◆ 주요센터 :교통안전공단, 국가교통DB(viewt.ktdb.go.kr)(교통연), 국가교통정보센터 등</p>	

⑨ 재난·안전	사회안전 분야에서 데이터 수집 및 가공 분석 플랫폼을 조성하여 데이터 기반 재난·안전 업무 및 대국민 서비스 강화
<ul style="list-style-type: none"> ○ 기존 재난관리자원 시스템을 위치데이터 생산·활용 GIS기반으로 고도화하고, 생활안전지도 확충*을 통해 데이터 개방 확대(행안부) * (현행) 교통, 재난, 치안, 맞춤형안전 → (추가) 시설, 산업, 보건, 사고안전 ○ 항공 안전 위해요소를 사전 차단할 수 있도록 항공안전데이터 수집·분석 플랫폼 구축(’18~’22, 국토부) ○ 기상·하천 등 실측자료 활용 통합홍수정보시스템 구축(~’22) 및 지역별 가뭄 취약지도 구축(~’18, 환경부) <p>◆ 주요센터 : 재난관리자원 공동활용시스템(행안부), 국가동물방역통합시스템(kahis.go.kr)(농림축산검역본부), 생활안전지도(국립재난안전연)(safemap.go.kr) 등</p>	

⑩ 환경	국민체감 환경 질 개선 및 환경오염 피해 예방을 위해 환경정보 데이터 수집·생성 관리체계 구축
<ul style="list-style-type: none"> ○ 기존 정보시스템*에 분산 운영되던 환경 데이터**를 통합하여 개방하는 환경정보 융합 플랫폼 구축(한국환경공단) * 대기오염, 수질오염, 도로먼지, 자동차 배출가스, 소음정보, 상하수도, 폐기물 관리, 자동차·전자제품 재활용 등 분야별 정보시스템 운영중 ** 대기, 물, 상하수도, 환경보건, 자연보전, 자원 재활용, 자원순환 등 40여개 분야 ○ 건물 에너지 사용량, 온실가스 배출정보 등을 수집·분석하는 ‘건물에너지 통합 진단 플랫폼’을 구축·확대*하여 에너지 소비 절감 및 관련 정책 수립 지원(~’21, 한국에너지공단) * 지자체 산하 공공기관(’17) → 전 공공기관(’19) → 국내 중대형 건물(’21) ○ 민간 분야에서 국가 대기질 및 IoT 기반 공기측정 데이터, 기상 데이터, 각종 환경 데이터 등을 실시간 수집·유통 제공하는 미세먼지 빅데이터 센터 구축 및 운영(’18~, 케이웨더) * 미세먼지 사전 예보, 영유아 및 어린이, 노령층 등 취약계층 행동 가이드, 살수차 운행 경로 설정 등 신규 서비스 개발 및 사회현안 해결 등에 활용 <p>◆ 주요센터 : 한국에너지공단, 한국환경공단, 기상청 등</p>	

1. 추진배경

- 4차 산업혁명 시대에는 **데이터가 국가와 기업의 경쟁원천**
 - 초연결 지능화가 도래하면서 사람과 기기 등에 의해 생성된 **대량의 데이터가 모든 산업의 발전과 새로운 가치창출의 촉매**로 작용
 - * 세계 데이터량(IDC, '17): ('16) 16ZB → ('25) 163ZB, 10배 증가 전망
 - **대규모 데이터를 보유하고 활용을 잘하는 기업이 시장을 주도**
 - * '17년 세계 시장 상위 5개 기업이 모두 데이터 기업 : 애플, 구글, MS, 아마존, 페이스북(시총기준)
 - 선진국은 **데이터 기반의 혁신성장을 이끄는 데이터 경제 활성화**에 역점
 - * (美) 빅데이터 R&D 전략('16) / (EU) 데이터경제 육성 전략('17) / (中) 빅데이터 발전계획('17) 등
- ➔ 국내 **데이터 생태계**를 되돌아보고 **선도국가로 도약하는 전략** 필요

2. 그간 데이터 정책의 한계

- **엄격한 개인정보 규제**로 인해 **데이터의 활용이 위축된 상황**으로 데이터 거래, 산업적 활용 등이 더디게 진행
 - * 우리나라의 정보제공 동의제도 등 규제는 OECD 주요국 중 가장 높은 수준(애널리시스 메이슨社, '14)
- 4차 산업혁명 시대 新제품·서비스 개발에 필수인 자율차·스마트시티 등 영역별 **실제 데이터*(Real Data)**와 **AI학습용데이터 구축이 미흡**
 - * (예) 자율주행차 : 정밀지도 데이터, 도로주행 센서·카메라 데이터, 교통사고 데이터 등
- 공공데이터가 일부 개방되었으나, **중소·스타트업이 필요로 하는 데이터는 부족하고 데이터 갱신·품질 수준이 떨어져 활용도 저조**, 공공·민간의 **클라우드 컴퓨팅 이용이 낮아** 빅데이터의 저장·관리가 비효율적
 - * 한국 2.5만개, 미국 23.3만개, 영국 4.4만개('18.3) / 정확성 결여, 불일치, 중복·누락 등 발생
 - * 우리나라의 클라우드 이용률 12.9%로 OECD 33개국 중 27위('17)
- 빅데이터 R&D와 전문인력 양성에 지속적인 투자에도 불구하고, **글로벌 수준의 기술·기업이 전무하고 전문인력도 가뭄 수준**
 - * 글로벌 100대 빅데이터기업에 한국기업 전무(美CPN社발표, '17) / 향후 3년간 1.3만명 필요

3. 비전 및 추진전략

비전

데이터를 가장 안전하게 잘 쓰는 나라

기본방향

데이터 경제 활성화로
혁신성장과 삶의 질 향상

데이터 보호와 활용의 균형,
글로벌 스탠다드 지향



추진전략 및 과제



4. 추진과제

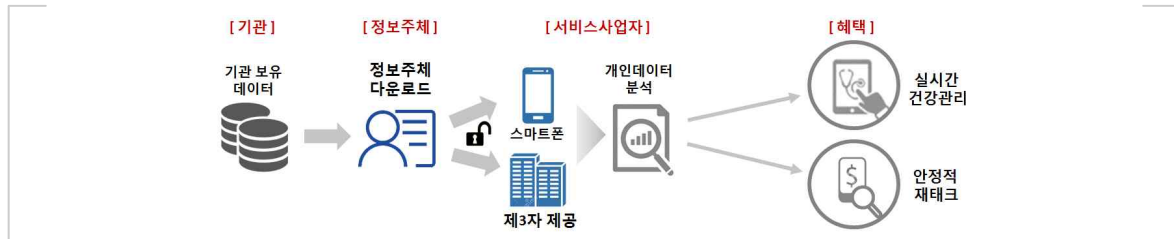
전략① 데이터 이용제도 패러다임 전환

◇ 글로벌 수준에 부합하도록 ▲정보주체인 개인이 자기정보를 스스로 통제·활용하는 패러다임으로 전환, ▲개인 데이터의 안전한 활용 전방위 지원(법제화·안심존구축·신기술적용·EU GDPR대응) 추진

① 데이터 이동권 확립 : 국민의 데이터 주권 찾기(MyData)

- 정보주체가 기관으로부터 자기 정보를 직접 내려받아 이용하거나 제3자 제공을 허용하는 방식으로 데이터 활용체계를 정보주체 중심으로 전환(예: 美 블루버튼)

➔ MyData 先도입은 개인정보 관련 법 개정 없이도 시행 가능하고, 비식별처리로 인해 데이터 활용가치가 낮아지는 문제도 극복 가능



- 의료금융통신 등 분야에서 대규모 시범사업 실시('18~)를 통해 국민 체감을 높여 건강증진·재택·통신비절감 등 실질적 효과를 토대로 先인식전환 후 제도화 추진
- * [의료](5대병원) 건강검진결과 스마트폰 다운로드>실시간 건강관리 [금융](100만고객) 계좌거래, 카드구매 내역 Open API 형태로 제공>빠른 자산 통합조회 [통신](20만고객) 음성·데이터 사용량 다운로드>맞춤형 요금제 추천

② 개인정보 안전한 활용 촉진

- 국민적 신뢰 기반 제도화를 위해 개인정보 해커톤의 사회적 합의 결과*를 바탕으로 개인정보보호법·위치정보법 등 법 개정 추진('18)

* 개인정보 범위 명확화, 비식별조치 근거인 가명·익명정보 개념 정립 등

- 데이터 자체의 반출은 안되나 데이터 분석 및 AI개발은 가능한, 보안환경을 갖춘 데이터 안심존 제공('19)

* 이용자가 물리적 보안 공간 또는 보안네트워크 원격시스템을 통해 다양한 분석SW(오픈소스 등)를 활용, 저장된 데이터(표본DB/맞춤형DB)를 분석해 볼 수 있는 온·오프라인 샌드박스 개념 형태

- 데이터 보안성이 높은 블록체인, 동형암호 등 신기술 적용실증 추진('18~)

- EU GDPR 시행('18.5.25)에 따른 우리기업 피해 최소화를 위해 대응 가이드라인 마련 및 적정성 평가* 승인 추진('18)

* 적정성 평가 미 승인 시, 개별 기업 단위로 개인정보 역외이전 승인·보호조치가 필요하여 기업 비용 증가

전략② 데이터 가치사슬 쏠주기 혁신

◇ **데이터 구축·개방**(4차 산업혁명 핵심데이터 등, ~'22) → **저장·유통**(데이터 거래의 비즈니스화, '18~) → **분석·활용**(산업·사회 혁신 활용, '18~) 등 **순과정에 걸쳐 실제데이터 기반** 영역별(의료·교통 등) **국가 빅데이터 지원체계** 마련

① **(구축·개방)** 4차 산업혁명의 핵심 기반인 산업별 **실제데이터**, **AI 학습 데이터**를 전방위 구축하고, **공공·민간 데이터의 획기적 개방**을 추진

- 모든 공공데이터는 **원시 데이터(raw data)** 형태로 **최대한 모으고**, 이중 민간 수요가 높은 데이터를 **국가중점데이터**로 선정('17. 48개 → '22. 128개 분야), **조기 개방 확대**
- **산업별 원시 데이터**의 풍부한 수집·생성을 위해 **빅데이터 전문센터**를 육성하고 각종 **빅데이터센터간 협력 네트워크 확대**('18~)



- 국내 인공지능(AI) 산업의 비약적 발전을 위해 **범용**(이미지·상식 등)·**전문**(법률·특허·의료 등) **분야 AI데이터셋**을 수요 중심으로 **단계적 구축보급**(~'22)
- * [범용데이터셋] '17. 688만건 → '22. 1.1억건 [전문데이터셋] '17. 211만건 → '22. 4,800만건
- 공공데이터 **전수조사** 및 **원칙적 개방**(국가안보·개인정보 제외), **공공부문 핵심 데이터**(통계·연구·참조표준 등) **구축개방 전면화, 품질 고도화**(오픈포맷 3단계 ↑)



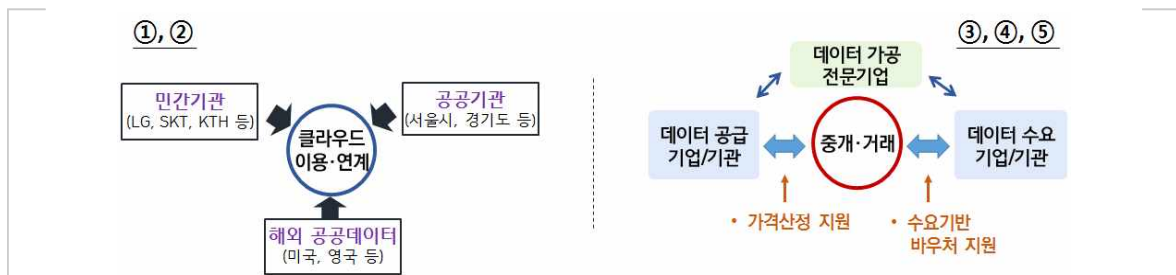
② **(저장·유통)** 데이터의 효율적인 저장·관리를 위한 **클라우드 본격 확산** 및 데이터 유통 촉진을 위해 민간·공공을 연계한 **개방형 데이터 기반 구축**

- 민간 클라우드서비스 이용기관을 공공기관에 한정하지 않고 **지자체 등으로 확대**(부처협의로 관련 제도 개선 추진)
- 스마트시티, 스마트공장, 창업 등에 클라우드 서비스를 접목하는 **All@Cloud 확산**, **중소·벤처기업 1만개 이상에 클라우드 도입 지원**(~'22)

○ 개방형 데이터 거래 체계 구축 등을 통한 양질의 데이터 유통 촉진

- 누구나 데이터를 한 곳에서 쉽고 빠르게 등록·검색·거래할 수 있도록 민간·공공을 연계한 데이터 거래 기반구축('18~)

①민간·공공 데이터 포털을 연계, 개방형(CKAN)으로 고도화 → ②국내·외 주요 데이터 플랫폼 연계 → ③데이터 가공 전문기업과 수요기업 매칭 → ④초기시장 활성화를 위해 중소·스타트업에 데이터 바우처(구매·가공비용 등 지원) 제공 → ⑤바우처를 통해 가공된 데이터는 상품으로 다시 등록·공유



- 데이터상품(날씨, 지도, 기업정보 등) 보유기업들이 협력하여 데이터를 공유·판매하는 **데이터 거래소**를 **설립·운영**시 **초기단계 지원**(‘18~)

* 공공데이터 가공, 오픈소스 분석SW 자원, 비식별처리 솔루션 등 연계 지원

- 데이터 거래의 **투명성 확보** 및 **공정경쟁 환경 조성**, **지재권 보호·활용 촉진** 등을 위한 **제도 연구·정비**(‘18~)

③ (분석·활용) 데이터를 활용한 산업적·사회적 가치의 본격 창출을 통해 산업 전반에 성장활력을 제고하고 국민 삶의 질 향상에 기여

- 전통 중소기업의 혁신 돌파구 마련을 위해 新제품 개발·맞춤형 홍보 등에 데이터를 활용하도록 **빅데이터 분석 전문기업**을 **매칭** 지원(500개사, ~'22)

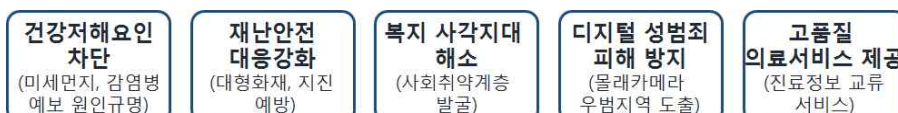
* (예) 티젠(茶 생산업체): 북미 소비자 맞춤형 茶 출시 → 해외시장 진출 한계 극복(30만불 수출, '17)

- 바이오, 미래소재, 대형연구장비 등 **연구데이터** 집약형 분야에서 **산업적 활용 촉진**을 위해 **선도 프로젝트*** 실시(‘18~)

* [바이오] 화합물 데이터(약 50만건) 활용 신약 개발 플랫폼 구축(시간·비용 1/3↓)

- 데이터기반 의사결정 지원 ‘**공공빅데이터센터**’ 설치, 미세먼지, 재난 안전 등 국민적 관심이 큰 사회문제 해결을 위해 **빅데이터 플래그십 프로젝트** 투자를 대폭 확대(‘18~)

* 빅데이터 플래그십 프로젝트 **투자확대** 목표 : (‘18) 8건 → (‘19) 20건 → (‘22) 25건

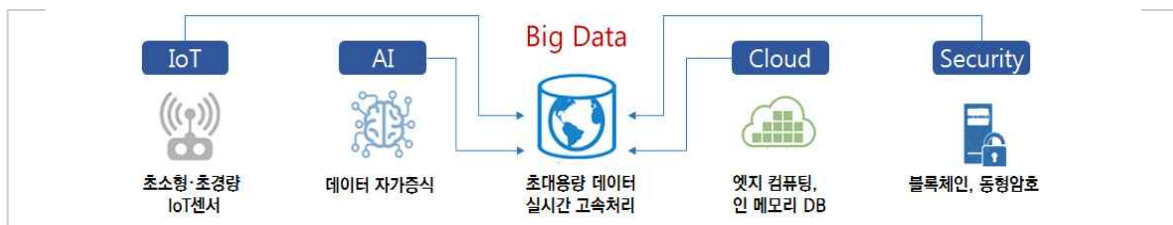


전략③ 글로벌 데이터산업 육성기반 조성

◇ 빅데이터 산업이 4차 산업혁명을 선도하는 **핵심 성장주체**로 거듭날 수 있도록 **4차산업혁명기술 융합, 분석 전문인력 양성, 컴퓨팅파워기반 기업성장 인프라 지원** 등을 아우르는 **역동적 산업 생태계 조성**

① 빅데이터 관련 선도기술 조기 확보 : 선진국 대비 90% 이상 수준

- 빅데이터 기술발전(분석→인공지능→초연결지능화)에 맞춘 **핵심 원천기술** 개발을 추진하고 4차 산업혁명 **요소기술(AIoT-클라우드정보보호)**과의 **융합 선도기술** 개발 가속화('18~'22)



② 미래수요 대응 전문인력 확충 : 청년고급인재, 실무인력 중심 5만명 양성

- 청년일자리 연계, 산업 수요기반 전문인력 양성을 위해 **빅데이터 전문교육 프로그램**을 지원하고 **데이터 분석 국가기술자격제도** 신설·운영
- 데이터 분석 고급인재 양성을 위해 **대학의 전공, 연구센터 운영 확대, 데이터 인재발굴 플랫폼** 구축·운영

* [과기정통부] 빅데이터 청년인재 일자리연계 사업, 데이터 아카데미, 교수·시간강사 연수 등
[고용부] 4차 산업혁명 선도인력 양성사업, 국가인적자원개발 컨소시엄 등

③ 빅데이터 전문기업 성장 지원 : 컴퓨팅자원, 맞춤형 사업 등을 통한 100개社 육성

- 글로벌 수준의 데이터 강소기업 육성을 위해 **컴퓨팅 자원 데이터셋** 등이 집적된 **판교 글로벌 ICT 혁신 클러스터** 고도화 및 **맞춤형 성장 지원** 강화

