

# III

## SW, 기반SW · 컴퓨팅 빅데이터





# 목차

## 빅데이터



### I. 표준화 개요

1.1. 기술 개요 .....	197
1.2. 표준화 비전 및 기대효과 .....	199
1.3. 표준화 추진체계 .....	201
1.4. 중점 표준화 항목 .....	202



### II. 국내외 현황분석

2.1. 연도별 주요 현황 및 이슈 .....	204
2.2. 정책 현황 및 전망 .....	205
2.3. 시장 현황 및 전망 .....	206
2.4. 기술개발 현황 및 전망 .....	208
2.5. IPR 현황 및 전망 .....	213
2.6. 표준화 현황 및 전망 .....	216
2.7. 오픈소스 현황 및 전망 .....	223



### III. 국내외 표준화 추진전략

3.1. 표준화 SWOT 분석 .....	226
3.2. 중점 표준화 항목별 국내외 추진전략 .....	227
3.3. 오픈소스 국내외 추진전략 .....	244
3.4. 중기(3개년) 및 장기(10개년) 표준화 계획 .....	245



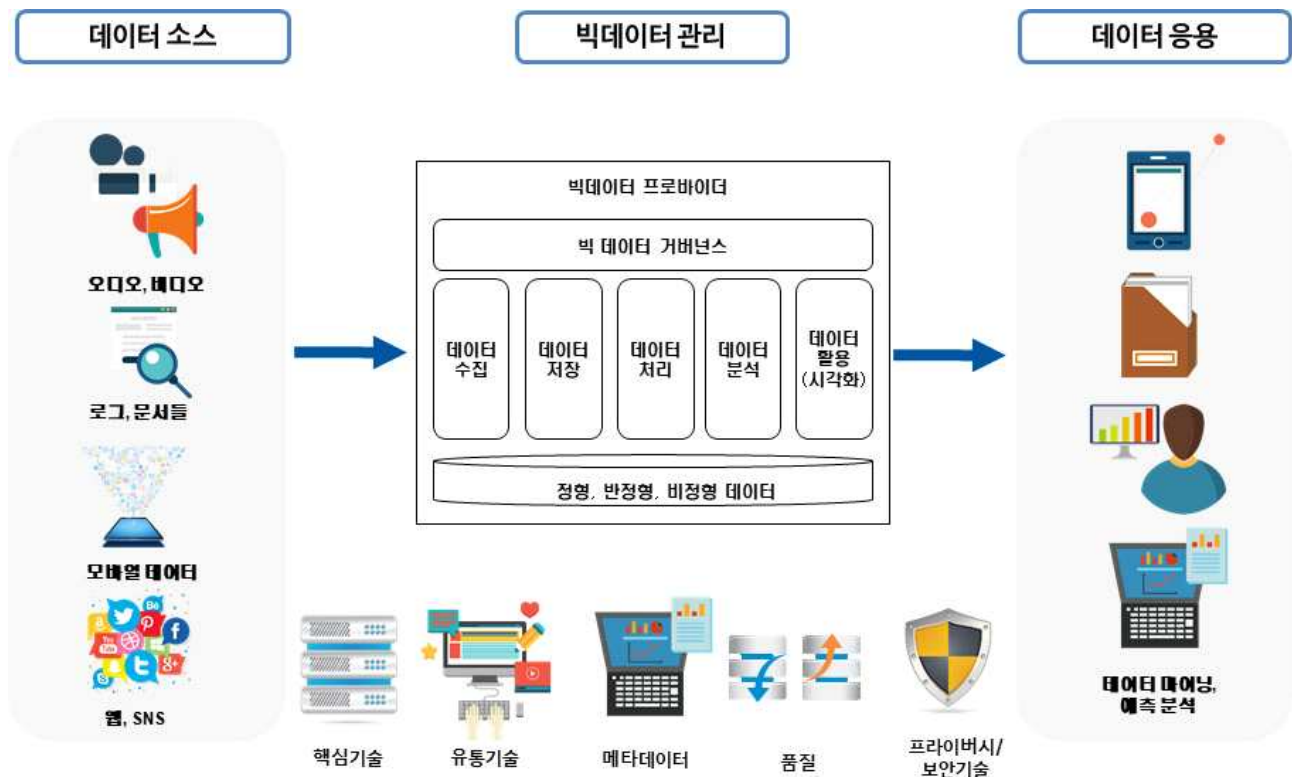
[작성위원] .....	247
[참고문헌] .....	248
[약어] .....	250



## I. 표준화 개요

### 1.1. 기술 개요

빅데이터는 기존의 데이터 아키텍처로 효과적으로 처리하기 어려운 데이터를 지칭하는 용어로서, 대용량 데이터(volume), 다양한 분야별 데이터 유형의 다양성(variety), 데이터 생성 시간의 가속화(velocity) 및 빠른 변화(variability) 등 특성을 지닌 데이터를 효과적으로 처리하기 위한 기술



<빅데이터 기술의 개요도>

- 빅데이터는 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고, 데이터의 초고속 수집, 발굴, 분석을 지원하도록 고안된 차세대 기술 및 아키텍처로서, 당초 수십~수천 테라바이트에 달하는 거대한 데이터 집합 자체만을 지칭하였으나, 점차 관련 도구, 플랫폼, 분석기법까지 포괄하는 용어로 변화

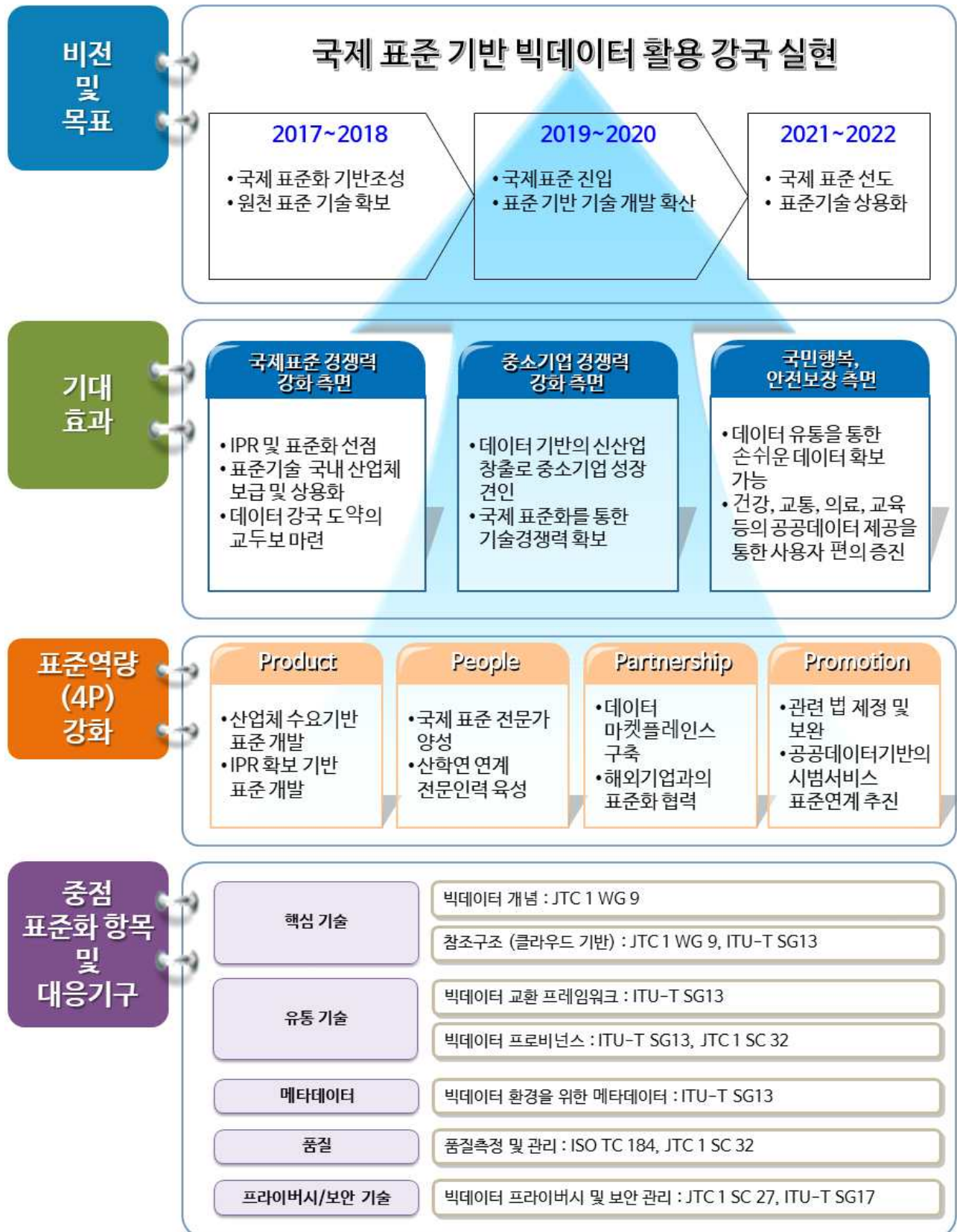
※ 매킨지 세계 연구소(2011), 삼성경제연구소 (2011), 가트너(2012), 오라클(2012) 등에서 빅데이터에 대한 정의를 다양하게 내리고 있으며, 현재 표준화 기구에서 빅데이터에 대한 정의를 개발 중. 이들의 공통된 정의를 규정짓는 특징은 양(Volume), 속도(Velocity), 다양성(Variety), 가치(Value) 등임

- 빅데이터는 “새로운 가치 추출을 위해 기존의 기술 또는 기법으로 처리하기 어려운 특징들(규모, 신속성, 가변성, 다양성, 진정성)을 갖는 데이터 모음”으로 정의

- 규모 (Volume) : 전통적으로 데이터 분석 서비스에 필요한 데이터의 규모는 테라바이트(terabyte) 규모가 통용되었으나 최근 테라바이트에서 페타바이트(petabyte)까지 그 규모가 커짐으로 인해 기존의 정보 기술 구조로는 효과적인 처리가 어렵게 됨. 이에 따라 수평적으로 연계 가능한 자원들간의 확장과 분산 질의 처리 방법들로 비 관계형 데이터 모델, NoSQL, 빅 테이블, 그래프 모델 등이 사용되고 있음
- 다양성 (Variety) : 기존의 데이터 분석에는 개별 도메인별로 구조화된 단일 데이터 유형을 사용. 하지만 최근 데이터 응용들은 정형, 반정형, 비정형 등 다양한 형태의 데이터들을 사용하고 있으며, 또한 이들 데이터들을 서로 조합함으로써 새로운 지식을 추출하는데 활용. 따라서 반정형/비정형의 데이터들을 보다 효과적으로 처리하고, 다양한 데이터 유형들을 여러 데이터 저장소에 효과적으로 저장하고 연계 활용하기 위한 방법들이 사용됨
- 신속성 (Velocity) : 신속성은 데이터가 생성되어, 저장, 분석, 시각화되는 과정에 필요한 시간으로 표현. 기존에는 트랜잭션 처리와 분석을 분리하여, 미리 저장된 데이터를 일괄 처리 방식으로 분석하여 사용. 그러나 실시간, 스트리밍 데이터 기반의 분석 요구가 증가하고 있으며, 이를 지원하기 위한 방법들이 모색되고 있음. 이는 기존에 이미 연구가 이루어지고 있다는 점에서 새로운 접근은 아니며, 빅데이터 관점에서는 이러한 데이터들을 효과적으로 다루기 위한 수평적 확장성에 초점이 맞추어져 있음
- 가변성 (Variability) : 가변성은 데이터의 포맷/구조, 의미, 품질 등과 같이 응용 또는 분석 유형에 따라 변하는 데이터의 특징을 의미. 이는 빅데이터의 다른 특징들이 변하는 것을 의미하며, 아키텍처, 인터페이스, 처리 알고리즘, 데이터 융합, 저장, 데이터 사용 등의 요인으로 인하여 발생. 예를 들어 빅데이터 아키텍처를 채용한 시스템일 경우에도 데이터를 용도에 따라 재가공하여 사용하거나 외부의 다양한 데이터와 융합하여 분석할 필요가 있으며, 따라서 데이터의 변화에 대한 신속한 대응이 필요
- 진정성 (Veracity) : 진정성 데이터에 대한 진실성, 사용 가능성, 왜곡 정도 등 데이터의 품질을 나타내는 것으로, 데이터의 융복합에 따른 원시 데이터에 대한 추적, 처리 방법 등에 대한 투명성을 보장하기 위한 기술적 장치가 필요

## 1.2. 표준화 비전 및 기대효과

### ○ 표준화 비전



## ○ 표준화 목표

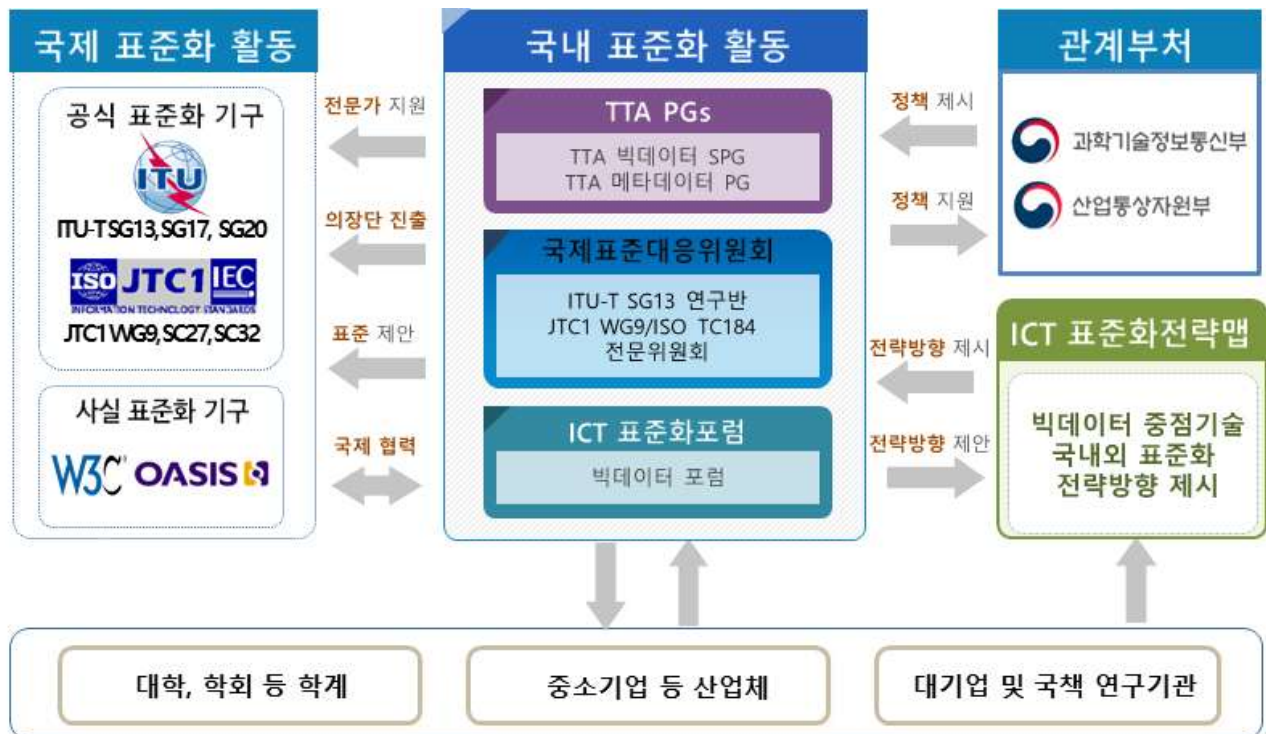
- 국내 빅데이터 관련 산·학·연은 지속적인 주도권 및 경쟁력 확보를 유지하기 위하여 다음과 같은 표준화 목표를 설정
  - 단기적으로 (2018년경까지), 국내 민간 포럼, 국내 R&D 과제들과 협력하여 표준 기획, 표준 과제 발굴, 표준(안) 개발을 통해 국내 표준 기반 조성 및 국제 표준화를 추진하기 위한 JTC 1 및 ITU-T 등의 빅데이터 관련 표준화 그룹의 의장단 및 에디터십 확보 추진
  - 중기적으로 (2020년경까지), 빅데이터 유통 활성화 기술 개발에 필수적인 국제 표준 기고를 통해 국가 기술 경쟁력을 확보하며, 국제 표준화를 선점하여 국제 표준을 선도
    - ◆ 빅데이터는 ITU-T SG13 및 ISO/IEC JTC1 WG 9에서 빅데이터에 대한 신규 표준 개발 추진
    - ◆ 국내 클라우드 및 빅데이터 기술개발 사업, 시범 사업 등의 국가 주도의 사업의 표준화 연계 노력
    - ◆ 표준특허 창출을 위한 클라우드 및 빅데이터 국제 표준 출원 증대 노력
  - 장기적으로 (2022년경까지), 2단계 활동으로 개발된 표준 기술에 대한 상용화를 추진함과 동시에 국제 표준 기고를 통해 국제 표준화를 선점하여 국제 표준을 선도

## ○ 표준화 기대효과

- 국제표준 경쟁력 강화 측면
  - 빅데이터의 IPR 및 표준화를 선점하고 국내 산업체 보급 및 상용화를 통하여 도메인별 클라우드 서비스를 포함한 SW 서비스 및 플랫폼 분야의 세계시장 경쟁력을 제고시키며, 차세대 SW 산업 강국 도약의 교두보 마련
  - 빅데이터의 수요자 중심의 기술 표준화 추진을 위하여 기술·제품에 대한 개방형 표준규격 개발 및 제품 간 (HW, SW 솔루션 등) 호환성 제공으로 국내 빅데이터 활성화와 국제 표준화를 통한 기술경쟁력을 확보
- 중소기업 경쟁력 강화 측면
  - 글로벌 기업의 기술·서비스 독점 및 종속성 극복을 통한 시장 공정경쟁 환경 마련 및 이용자 편익 증진
  - 빅데이터 산업을 통한 신생 기업 활성화로 고용 창출 효과
- 국민행복·안전보장 측면
  - 국내의 빅데이터 인력수요가 증가할 것으로 예상되며 빅데이터 일자리 창출 효과



### 1.3. 표준화 추진체계



#### ○ ICT 표준화전략맵

- 표준화전략맵의 표준화 전략방향에 따라 TTA 빅데이터 SPG를 통해 단체표준을 개발과 ITU-T 연구반과 JTC1 전문위원회의 국제 표준화 전략 방향을 제시

#### ○ 국내 표준화 활동 체계

- 국내 빅데이터 포럼에서 산학연 의견수렴하여 포럼표준을 제정하고, TTA 빅데이터 SPG를 통해 단체표준을 개발

#### ○ 국제 표준화 활동 체계

- ITU-T SG13, SG17, SG20와 JTC 1 WG9, SC27, SC32에서 주도적으로 표준을 개발함
- 사실표준화기구인 W3C, OASIS 를 통해 대응

## 1.4. 중점 표준화 항목

### ○ 중점 표준화 항목 범위의 설정

- 핵심표준 기술은 빅데이터 개념 및 정의, 빅데이터 참조 구조, 클라우드 기반 빅데이터 서비스 기능 및 아키텍처 3개 항목을 중점 표준화 항목으로 설정
- 유통 기술은 빅데이터 교환을 위한 프레임워크 및 요구사항, 빅데이터 프로비넌스 요구사항 2개 항목을 중점 표준화 항목으로 설정
- 메타데이터 기술은 빅데이터 환경을 위한 메타데이터 1개 항목을 중점 표준화 항목으로 설정
- 품질 기술은 빅데이터 품질측정 및 관리 모델 1개 항목을 중점 표준화 항목으로 설정
- 프라이버시/보안기술은 빅데이터 프라이버시 및 보안 관리 1개 항목을 중점 표준화 항목으로 설정

중점 표준화 항목		표준화 내용	Target SDOs	전략 목표
핵심 표준 기술	빅데이터 개념 및 정의	- 빅데이터의 일반 개념과 빅데이터 분야에서 공통적으로 사용되고 있는 용어들을 식별하고 그 의미를 정의	JTC1 WG9 ITU-T SG13	전략적 수용
	빅데이터 참조 구조	- 빅데이터 생태계를 구성하는 각 요소 (데이터 제공자, 빅데이터 서비스 제공자, 빅데이터 소비자)들 대한 역할 정의 및 이들 간의 관계를 정의하는 표준	JTC1 WG9, ITU-T SG13	적극 공략
	클라우드 기반 빅데이터 서비스 기능 및 아키텍처	- 클라우드 기반의 빅데이터 제공을 위한 기본 요소 기능 정의 - 클라우드기반의 빅데이터 제공에 필요한 클라우드구조(CCRA)의 확장 정의 - Big data as a Service(BDaaS) 구성요소들 간의 인터페이스 정의 표준	ITU-T SG13	적극 공략
유통 기술	빅데이터 교환을 위한 프레임워크 및 요구사항	- 빅데이터 환경에서 서비스 제공자가 데이터들을 수집하여(import) 가공된 데이터 형태로 서비스(export)하기 위한 프레임워크 및 기능 요구사항을 정의	ITU-T SG13	차세대 공략
	빅데이터 프로비넌스 요구사항	- 빅데이터 환경에서 데이터의 융합 및 변경 등에 대한 이력 및 출처에 대한 추적을 지원하기 위한 기능 요구사항을 정의	ITU-T SG13	차세대 공략
메타 데이터	빅데이터 환경을 위한 메타데이터	- 빅데이터 환경에서 저장되는 도메인별 메타데이터의 상호운용성을 확보하기 위한 개념, 등록, 관리, 거래, 활용에 대한 표준	JTC1 WG9, ITU-T SG13	차세대 항목
품질	빅데이터 품질측정 및 관리 모델	- 빅데이터 품질 관리 요구사항 정의 표준 - 빅데이터처리 과정에 따른 이력관리 및 품질요소 측정표준 - 활용 용도기반의 빅데이터 품질요소 정의 표준	JTC1 SC32 WG2, ISO TC184 SC4 WG13	적극 공략
프라이버시/보안 기술	빅데이터 프라이버시 및 보안 관리	- 빅데이터 참조 모델에서의 보안 역할 및 책임, 보안 및 프라이버시 패브릭 - 모바일 인터넷서비스에서의 보안챌린지, 보안위협 및 보안 요구사항 - 서비스로서의 빅데이터 환경에서의 보안챌린지, 역할 및 책임, 보안대책	JTC1 SC27, ITU-T SG17	차세대 공략

## ○ 추진경과

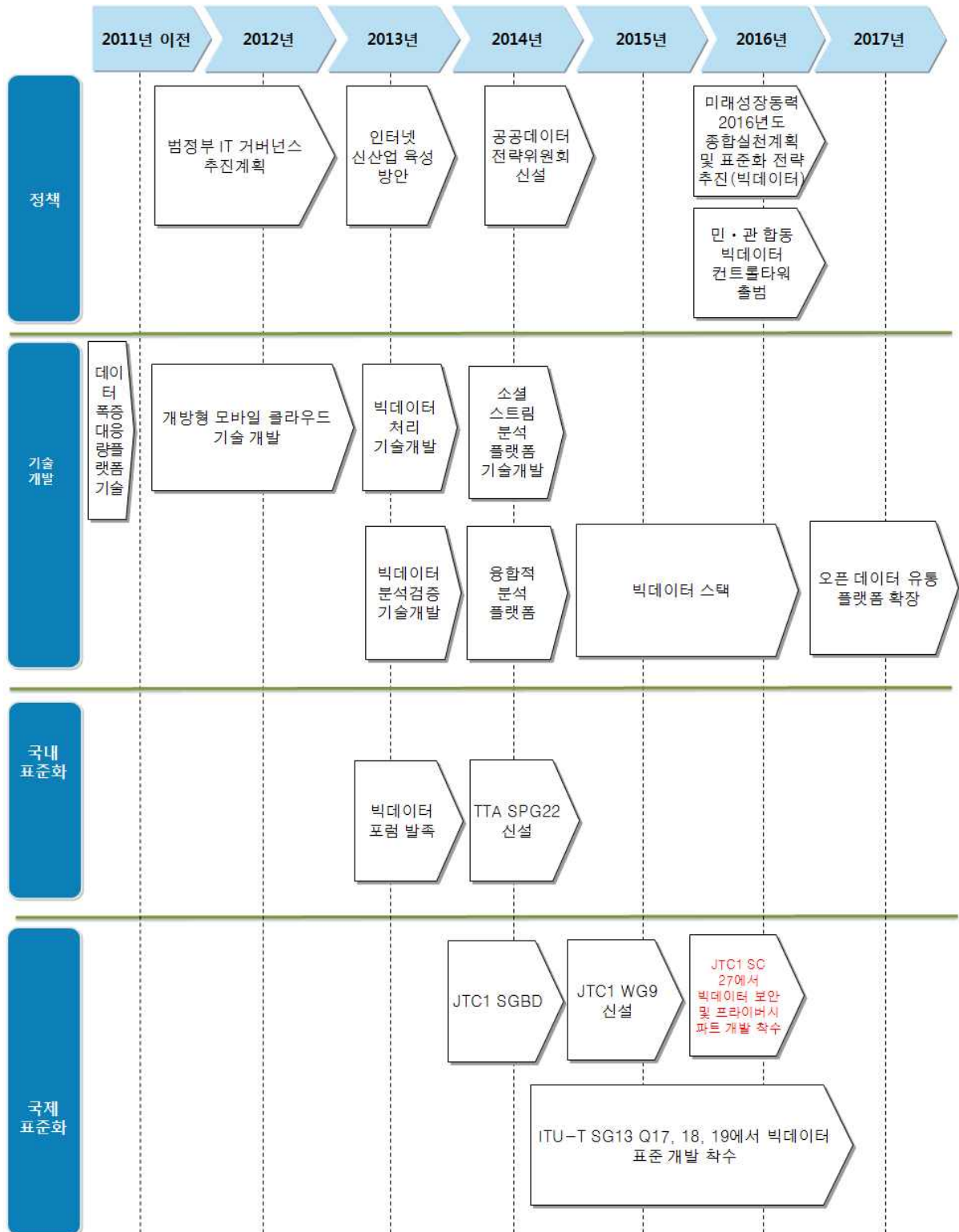
- Ver.2016(2015년)에서는 표준화 활동이 중단되거나 없는 표준화 항목(빅데이터 분석활용 가이드) 중심으로 소폭 제외되었으며, 빅데이터 분야의 신규 표준화 항목(빅데이터 교환을 위한 프레임워크 및 요구사항)을 추가
- Ver.2017(2016년)에서는 신규 표준화 항목으로 빅데이터의 연역 관리를 위한 빅데이터 프로비넌스가 추가
- Ver.2018(2017년)에서는 Ver.2017에서 클라우드컴퓨팅과 빅데이터를 분리하였으며, 항목 변경은 없으나, 빅데이터 프라이버시 및 보안 관리, 빅데이터 환경을 위한 메타데이터 등의 항목은 보다 구체화된 이름으로 변경되었음

&lt;버전별 표준화 항목 비교표&gt;

구분	Ver.2016	Ver.2017	Ver.2018
핵심표준기술	빅데이터 개념 및 정의	빅데이터 개념 및 정의	빅데이터 개념 및 정의
	빅데이터 참조 모델	빅데이터 참조 모델	빅데이터 참조 구조
	-	클라우드 기반 빅데이터 서비스 기능 및 아키텍처	클라우드 기반 빅데이터 서비스 기능 및 아키텍처
유통기술	빅데이터 교환을 위한 프레임워크 및 요구사항	빅데이터 교환을 위한 프레임워크 및 요구사항	빅데이터 교환을 위한 프레임워크 및 요구사항
	-	빅데이터 프로비넌스 요구사항	빅데이터 프로비넌스 요구사항
메타데이터	빅데이터를 위한 메타데이터	빅데이터를 위한 메타데이터	빅데이터 환경을 위한 메타데이터
품질	빅데이터 품질 측정 및 관리 모델	빅데이터 품질 측정 및 관리 모델	빅데이터 품질측정 및 관리 모델
프라이버시/보안기술	빅데이터 보안 및 개인정보 보호를 위한 기능	빅데이터 보안 및 개인정보 보호를 위한 기능	빅데이터 프라이버시 및 보안 관리

## II. 국내외 현황분석

### 2.1. 연도별 주요 현황 및 이슈



## 2.2. 정책 현황 및 전망

구분	주요 현황
한국	<ul style="list-style-type: none"> <li>- 빅데이터 확산을 위한 39억 규모의 '2017년 빅데이터 플래그십' 프로젝트 12개 과제 선정해 시범사업 추진 [2017.6, 과학기술정보통신부]</li> <li>- 빅데이터산업 육성을 위한 K-ICT 전략 2016 발표 [2016.5, 구 미래부]</li> <li>- '빅데이터 산업 육성을 위한 2016년도 미래성장동력 종합실천계획' 발표 [2016.3, 구 미래부]</li> <li>- '빅데이터 개인정보보호 가이드라인'을 발표(2014.12.23., 방송통신위원회)</li> <li>- 2014년부터 TTA는 빅데이터 PG(SPG22)를 신설하고 빅데이터 플랫폼에 대한 요구사항 표준 표준개발을 착수하였으며, 빅데이터 참조 아키텍처, 데이터 서비스 개념, 정의 및 요구사항 등 관련 표준 개발을 추진 중[2014 ~ 현재]</li> <li>- '빅데이터산업 발전 전략' 수립 및 추진, '정부 3.0'으로 공공데이터 개방 추진 등 [2013, 구 미래부]</li> <li>- 사물인터넷, 클라우드, 빅데이터, 미래인터넷을 인터넷 신산업으로 지정하고 「인터넷 신산업 육성 방안」 발표[2013.6, 구 미래부]</li> </ul>
미국	<ul style="list-style-type: none"> <li>- 아마존, MS, 세일즈포스닷컴 등 기존 글로벌 기업들은 클라우드와 IoT, 빅데이터를 결합하여 제조업 등 다양한 산업에 적용 중</li> </ul>
일본	<ul style="list-style-type: none"> <li>- 일본 정부는 2012년 빅데이터 활용특별부회를 설립하여 빅데이터 활용을 위한 정책의 기본방향을 검토하여 '빅데이터 활용 방안'을 발표하여 진행 중</li> <li>- 2016년 일본 경제 균형발전에 빅데이터 활용하기 위하여 일본 경제산업성에서 지방 빅데이터 체계 구축사업 개시</li> <li>- 또한, 일본의 주요 기업들(예, 후지쯔 (의료분야), 로손 (신상품개발))에서 빅데이터 활용 서비스를 추진 중</li> </ul>
유럽	<ul style="list-style-type: none"> <li>- 유럽연합은 2015년부터 빅데이터에 대해 향후 5년 동안 산업 컨소시엄의 20억 유로와 공적 자금 5억 유로를 합쳐 총 25억 유로, 즉 3조 3,870억 원을 투자함</li> <li>- 2015년 1월부터 아토스(Atos), IBM, 노키아 솔루션앤네트웍스(Nokia Solutions and Networks, NSN), 오렌지(Orange), SAP, 지멘스(Siemens) 등의 기업들과 여러 연구기관들이 주도로 한 민관 협력(public-private partnership, PPP)을 통해 진행 중</li> </ul>
중국	<ul style="list-style-type: none"> <li>- 중국 정부는 빅데이터 산업을 13.5 계획(2016~2020년)에서 집중 육성 대상으로 지정, 핵심산업으로 육성 중임.</li> <li>- 이미 중국 정부는 '2013~2015년 정부 빅데이터 관련 프로젝트'를 수행하였고, 정책 지원 이외에도 정부 차원의 프로젝트를 통해 빅데이터 산업 발전을 추진 중</li> <li>- 특히, 국무원, 공업과 정보화부, 재정부 등에서도 관련 정책을 통해 빅데이터 기술 연구 개발 및 서비스 확대에 전력을 기울이고 있음</li> <li>- 또한, 지역 중심으로 빅데이터 산업 활성화를 추진 중이며, 한 예로 광둥성(廣東省)은 2014년 중국 최초의 빅데이터 관리기관인 '광둥성 빅데이터 관리국'을 설립하고, 빅데이터 응용 시범 작업을 진행한 후, 그 해 12월 '광둥성 빅데이터 발전 계획(2015-2020년)'을 발표하여 진행 중</li> </ul>

## 2.4. 기술개발 현황 및 전망

기술개발 수준	국내	<input type="checkbox"/> 기초연구 → <input type="checkbox"/> 실험 → <input checked="" type="checkbox"/> 시작품 → <input type="checkbox"/> 제품화 → <input type="checkbox"/> 사업화	국내외 격차	2년
	국외	<input type="checkbox"/> 기초연구 → <input type="checkbox"/> 실험 → <input type="checkbox"/> 시작품 → <input type="checkbox"/> 제품화 → <input checked="" type="checkbox"/> 사업화		

### 2.4.1. 국내 기술개발 현황 및 전망

- (빅데이터 개념 및 정의) 기존의 데이터 활용 사례들을 분석함으로써, 빅데이터 활용 우수사례 등을 발굴하고, 공공 빅데이터 활용 방안 모색에 집중. 가트너, 맥킨지 등 해외 기관들의 정의를 차용하고 있음
- (빅데이터 참조 구조) 빅데이터 분야에 공통적으로 활용가능한 포괄적인 플랫폼 보다는 개별 프로젝트 또는 빅데이터 저장기술, OLAP기반 빅데이터 분석 기술, 시각화 기술 등에 집중
  - (티맥스소프트) 빅데이터 처리 및 운용을위한 TIBERO 출시
  - (인피니플렉스) 빅데이터 환경에서 다양한 질의처리를 동시에 지원하는 InfiniFlux 출시
- (클라우드 기반 빅데이터 서비스 기능 및 아키텍처) 빅데이터 분석을 활용하고자 하는 업계의 수요가 지속적으로 상승하고 있음. 또한 클라우드를 활용하여 빅데이터를 분석 개발 중
  - (모비젠) 정부의 GCS(Global Creative SW)사업의 일환으로 퍼블릭 클라우드 기반의 초당 100만 레코드 이상 실시간 인덱싱을 보장하는 멀티테넌시 분산 빅데이터 분석 플랫폼 개발을 진행중에 있음. 빅데이터DB 및 분석 플랫폼을 상용 퍼블릭 클라우드 상에서 구현 및 관리할 수 있는 IoT 서비스 인프라의 개발이 목표임
  - (인브레인) 정부의 GCS(Global Creative SW)사업의 일환으로 클라우드 환경에서 빅데이터 기반 통합 데이터 분석 스위트 개발 중에 있음
  - (기타) 국내 빅데이터 관련 기업은 데이터랩을 제공하는 네이버를 비롯하여, 다음소프트, 솔트룩스, 그루터 등이 빅데이터 검색, 마이닝, 시맨틱 웹 관련 기업들이 주를 이루고 있음. 대부분 기업 제품의 경우 스토리지, 빅데이터 수집, 검색에서 빅데이터 분석에 이르는 하나의 플랫폼을 제공하고 있음
- (빅데이터 교환을 위한 프레임워크 및 요구사항) 빅데이터 유통을 위한 정부 주도의 시범사업 및 플랫폼 개발에 착수하였으며, 기존의 데이터 유통 사업자들을 수평적으로 연계하기 위한 기술개발을 진행 중
  - (데이터진흥원) 데이터스토어를 통해 공공 및 문화 데이터 유통
  - (ETRI) 웹기반의 데이터 카탈로깅 관련 표준인 DCAT 기반의 오픈 데이터 유통 플랫폼 기술 개발을 추진 중



- (빅데이터 프로비넌스 요구사항) 정부 R&D 계획을 중심으로 빅데이터에 대한 이력관리 및 품질관리를 지원하기 위한 기술개발을 기획하는 단계임
  - (그루터) 하둡 기반의 빅데이터 플랫폼을 통합 관리하는 솔루션으로 데이터 분석 전 과정에 대한 모니터링 및 이력 관리 기능을 제공하는 Cloumon 출시
- (빅데이터 환경을 위한 메타데이터) 빅데이터 환경을 위한 메타데이터에 대한 관심이 커지고 있으며, 공공 데이터 활성화를 위한 LOD(Linked Open Data) 적용이 추진되었으나 확산에는 한계를 보임. 따라서 빅데이터 환경에서 다양한 도메인 간의 데이터의 수집, 등록, 유통, 거래, 관리를 위한 메타데이터 표준 기술 개발이 필요한 상황임.
  - (데이터진흥원) 데이터 스토어를 개방형 플랫폼으로 전환하고, 2018년에는 데이터 거래소로 명칭을 바꿔 공식 거래소를 개장할 계획이며, 데이터 거래를 위한 기술 모델 표준을 기획
- (빅데이터 품질측정 및 관리모델) 대용량 데이터, 실시간 데이터 등 빅데이터에서 발생하는 오류를 자동으로 검출, 분류하는 기술 개발 중
  - (한국정보화진흥원) 공공정보 품질관리 매뉴얼(2014)을 개발하여 공공기관을 대상으로 공공데이터의 품질관리 개선과 데이터 품질 진단을 매년 실시 중
  - (한국데이터진흥원) 데이터품질관리 지침, 데이터 품질관리 성과 측정기법 등 품질관련 지침을 배포하여 데이터 품질관리 인증을 실시 중
  - (위세아이텍) 빅데이터 지능형 품질평가 솔루션 개발 중
  - (지티원) DQMiner를 통해 자동 데이터 프로파일링, 데이터 오디팅, 데이터 를 관리, 데이터 품질 분석, 데이터 품질 분석 결과 보고 등 제공
- (빅데이터 프라이버시 및 보안관리) 빅데이터 보안 기술은 빅데이터 이해당사자들의 보안 역할 및 책임과 빅데이터 운영을 위한 보안 및 프라이버시 측면의 다계층 기능 명세를 제공함. 또한 모바일, 서비스로서의 빅데이터 등 다양한 빅데이터 환경에서의 보안 위협 및 요구사항을 명세하고 빅데이터 분석 시 개인정보를 안전하게 처리하기 위한 비식별화 기법을 개발함
  - (국가보안연구소) 빅데이터 시스템의 보안 위협 및 보안 요구사항 분석
  - (효성인포) 빅데이터 기반 영상 보안 솔루션 출시

## &lt;국내 주요 사업자 서비스 동향&gt;

사업자	주요 현황																		
데이터진흥원	<ul style="list-style-type: none"><li>- 2017년 8월 현재 데이터스토어를 운영 중이며, 2018년에 개방형 플랫폼으로 전환하여 데이터거래소로 확장할 계획임.</li></ul>																		
엑셈(ex-em)	<ul style="list-style-type: none"><li>- IT 성능관리 및 빅데이터 플랫폼 전문 기업으로 R, MapReduce, Hive, Pig, YARN을 하나의 단일 환경에서 분석할 수 있도록 통합 환경을 제공하는 Flamingo 솔루션 (Big Data 운영 · 관리 · 분석 솔루션) 제공 중</li></ul>																		
솔트룩스 (Saltlux)	<ul style="list-style-type: none"><li>- 인공지능 기반의 빅데이터 분석과 시맨틱 검색 솔루션, 데이터 서비스 플랫폼 사업과 다국어 로컬리제이션을 포함한 언어/문서 서비스 사업을 수행</li><li>· 빅데이터 솔루션인 BigO는 자연언어처리(NLP)와 기계학습, 분산병렬 처리와 같은 핵심 기반 기술과 80억건 이상의 소셜 및 공공 빅데이터에 기반하여 빅데이터에 대한 시맨틱 검색/분석, 지능화뿐 아니라 플랫폼을 통해 IoT 센서와 생산 및 운영시스템 로그와 같은 스트림 빅데이터에 대한 실시간 분석, 예측 기능을 제공하는 플랫폼. 이를 기반으로 다양한 제품군을 출시(TORNADO, DISCOVERY, STORM, BLUEBOLT, LEA, RAINBOW)</li><li>· 인공지능 솔루션인 ADAMs는 현재 국내 최초로 상용화된 인공지능 플랫폼으로 대규모의 지식베이스를 내장하고 있으며, 50여개의 API를 제공</li></ul>																		
와이즈넷 (wisenut)	<ul style="list-style-type: none"><li>- 인공지능기반 빅데이터 분석&amp;수집, 검색솔루션 전문 기업으로 다양한 빅데이터 제품군을 보유</li></ul> <table><tr><th>구분</th><th>제품명</th><th>설명</th></tr><tr><td>지능정보 솔루션</td><td>WISE iDESK WISE i Chat WISE Shop Bot</td><td>o 인공지능경망 기술을 활용한 전문분야 콘텐츠/문서 추천, 민원 자동 대응/처리, 고객 정보 기반 상품 추천 기능 제공</td></tr><tr><td>수집 솔루션</td><td>WISE BICrawler WISE InfoFinder</td><td>o 빅데이터 수집과 언어분석 기술 기반, 소셜 데이터 분석에 최적화된 빅데이터 정보수집 o 영역/형태에 관계없이 원하는 정보를 자동으로 추출 및 필터링하여 선별하여 정보 수집</td></tr><tr><td>검색 솔루션</td><td>Search Formula-1 WISE Referee WISE MinAX</td><td>o 빅데이터 분석 / 대용량 통합검색 / 시맨틱 / 마이닝 / 중복문서처리 / 유사도 분석 등의 주요기술을 접목한 차세대 검색 기능제공</td></tr><tr><td>분석 솔루션</td><td>WISE TEA v2 WISE BIC Analyzer WISE Classifier WISE e-Discovery</td><td>o 다양한 비정형 빅데이터를 분석하여 주제(Topic), 의미분석, 카테고리 분류 등의 기능제공 o 기업내 비정형 빅데이터 외부유출 방지</td></tr><tr><td>서비스 솔루션</td><td>BuzzInsight BrandFair TweeTrend AD Plus</td><td>o 블로그/카페/SNS/뉴스 등에서 데이터를 수집하여 이슈 모니터링 분석, 브랜드 점유율 분석, 트위터 검색 기능 제공 o (AD Plus) 광고 서비스 기능</td></tr></table>	구분	제품명	설명	지능정보 솔루션	WISE iDESK WISE i Chat WISE Shop Bot	o 인공지능경망 기술을 활용한 전문분야 콘텐츠/문서 추천, 민원 자동 대응/처리, 고객 정보 기반 상품 추천 기능 제공	수집 솔루션	WISE BICrawler WISE InfoFinder	o 빅데이터 수집과 언어분석 기술 기반, 소셜 데이터 분석에 최적화된 빅데이터 정보수집 o 영역/형태에 관계없이 원하는 정보를 자동으로 추출 및 필터링하여 선별하여 정보 수집	검색 솔루션	Search Formula-1 WISE Referee WISE MinAX	o 빅데이터 분석 / 대용량 통합검색 / 시맨틱 / 마이닝 / 중복문서처리 / 유사도 분석 등의 주요기술을 접목한 차세대 검색 기능제공	분석 솔루션	WISE TEA v2 WISE BIC Analyzer WISE Classifier WISE e-Discovery	o 다양한 비정형 빅데이터를 분석하여 주제(Topic), 의미분석, 카테고리 분류 등의 기능제공 o 기업내 비정형 빅데이터 외부유출 방지	서비스 솔루션	BuzzInsight BrandFair TweeTrend AD Plus	o 블로그/카페/SNS/뉴스 등에서 데이터를 수집하여 이슈 모니터링 분석, 브랜드 점유율 분석, 트위터 검색 기능 제공 o (AD Plus) 광고 서비스 기능
구분	제품명	설명																	
지능정보 솔루션	WISE iDESK WISE i Chat WISE Shop Bot	o 인공지능경망 기술을 활용한 전문분야 콘텐츠/문서 추천, 민원 자동 대응/처리, 고객 정보 기반 상품 추천 기능 제공																	
수집 솔루션	WISE BICrawler WISE InfoFinder	o 빅데이터 수집과 언어분석 기술 기반, 소셜 데이터 분석에 최적화된 빅데이터 정보수집 o 영역/형태에 관계없이 원하는 정보를 자동으로 추출 및 필터링하여 선별하여 정보 수집																	
검색 솔루션	Search Formula-1 WISE Referee WISE MinAX	o 빅데이터 분석 / 대용량 통합검색 / 시맨틱 / 마이닝 / 중복문서처리 / 유사도 분석 등의 주요기술을 접목한 차세대 검색 기능제공																	
분석 솔루션	WISE TEA v2 WISE BIC Analyzer WISE Classifier WISE e-Discovery	o 다양한 비정형 빅데이터를 분석하여 주제(Topic), 의미분석, 카테고리 분류 등의 기능제공 o 기업내 비정형 빅데이터 외부유출 방지																	
서비스 솔루션	BuzzInsight BrandFair TweeTrend AD Plus	o 블로그/카페/SNS/뉴스 등에서 데이터를 수집하여 이슈 모니터링 분석, 브랜드 점유율 분석, 트위터 검색 기능 제공 o (AD Plus) 광고 서비스 기능																	



## 2.4.2. 국외 기술개발 현황 및 전망

- (빅데이터 개념 및 정의) 가트너, IDC, SAS, NIST 등 다양한 기관에서 각각 빅데이터에 대해 정의하고 있으며, 초기에는 데이터의 규모를 강조하였으나 최근에는 신규 가치 창출에 초점을 맞추고 있음
  - (구글, IBM, MS) 자사의 솔루션이 제공하는 기능들을 중심의 빅데이터 개념 정의
- (빅데이터 참조 구조) 빅데이터 분석 솔루션 제공을 위한 컨소시엄 또는 개별 사업체들의 통합 플랫폼에서 제공 가능한 기술 요소들을 중심으로한 빅데이터 참조 모델을 제시함
  - (구글) 빅데이터 시스템 환경에 필요한 컴포넌트들로 구성된 생태계 모델 제시
  - (HP) Apach Spark을 중심으로한 자사 소프트웨어 및 하드웨어 솔루션 컴포넌트로 구성된 참조 아키텍처 제시
  - (Oracle) 데이터 분석 중심의 빅데이터 처리기능을 강조한 개념 모델 제시
- (클라우드 기반 빅데이터 서비스 기능 및 아키텍처) 클라우드 서비스 제공 업체들은 빅데이터 데이터 저장소뿐만 아니라, 대용량 데이터의 분석 및 관리를 위한 기능들을 개발 제공하는 추세임
  - (트레저데이터) 클라우드 기반 빅데이터 관리 서비스 기업으로, 데이터 수집, 저장, 실시간 예측 분석, BI툴과의 연동을 제공 중
  - (구글) 클라우드 플랫폼의 빅데이터 분석 기술인 ‘클라우드 데이터플로우(Cloud Dataflow)’를 서비스 중. 페타 바이트급의 데이터 저장 및 분석용 클라우드 서비스 개발하여 제공 중
  - (아마존) EMR (Elastic MapReduce)를 개발하여 AWS 기반 대량의 데이터를 신속하고 비용 효율적으로 처리할 수 있도록 관리형 하둡 프레임워크를 제공. EMR은 Apache Spark, HBase, Presto 및 Flink와 같은 오픈 소스 프레임워크를 실행할 수 있도록 제공. 대화식 쿼리 서비스로 저장된 데이터를 분석하는 Amazon Athena 등 다양한 서비스를 제공 중
  - (페이스북) 대용량 빅데이터 분석처리 엔진 ‘프레스토’를 오픈소스 소프트웨어로 개발
- (빅데이터 교환을 위한 프레임워크 및 요구사항) 세계적으로도 공공분야에서 축적된 데이터들에 대한 개방 및 공유를 통해 데이터 비즈니스 분야의 활성화를 위한 노력들이 진행 중이며 이를 지원하기 위한 솔루션들이 개발 중
  - (BDEX) 빅데이터의 유통 및 거래를 지원하는 통합 플랫폼 출시

- (빅데이터 프로비던스 요구사항) 빅데이터에 대한 이력관리 기능을 지원하기 위하여 기존 DBMS 솔루션을 확장하거나 신규 플러그인 소프트웨어에 대한 개발을 진행 중
  - (TATA) 데스크톱 환경에서 빅데이터에 대한 이력추적 및 관리를 지원하는 TCS BDD 출시
- (빅데이터 환경을 위한 메타데이터) 빅데이터의 메타데이터의 중요성에 대한 인식이 확산되면서 데이터 전문회사, 오픈소스 진영, 표준화 단체에서 솔루션들을 제시하기 시작함
  - (Waterline Data, Collibra, Apache Foundation) 하둡에서 수집된 빅데이터를 자동으로 분류하고 메타데이터를 추가하는 솔루션을 소개하고 있으나, 아직 초기단계이며 주요 목표도 수집된 빅데이터를 자동화된 데이터 분류와 관리에 맞춰져 있음
  - (W3C) 웹 상의 데이터 카탈로그간의 상호운용성을 지원하기 위한 RDF 온톨로지로 DCAT(Digital Catalog Vocabulary)이 2014년에 표준화가 되었으며, 빅데이터에 적용 중
- (빅데이터 품질측정 및 관리모델) 빅데이터 중 비정형 데이터는 조직의 모든 데이터 중 70 ~ 80 % 이상을 차지하는 것으로 추산되며 이는 데이터 품질 도구 시장의 중요한 성장 영역으로 각광받고 있음
  - (인포메티카) 사물인터넷(IoT) 등 대용량 데이터 분석, 데이터 거버넌스 및 콘텐츠 데이터 분석, 기계학습, 알고리즘 및 예측 분석 적용 기술 개발
  - (IBM) Apache Kafka 오픈 소스 활용한 사물인터넷(IoT) 데이터 및 기계학습 지원 기술 개발
- (빅데이터 프라이버시 및 보안관리) 빅데이터 환경에서의 보안 및 개인정보 보호를 위한 기술이 개발되고 있으며 IBM, Syncsort 등은 자사의 솔루션에 보안 및 개인정보 보호를 위한 기능들을 추가하거나 보안 기능에 특화된 솔루션 들을 출시
  - (IBM) 빅데이터 환경에서 위험요소를 자동으로 식별하여 차단하는 IBM Security Intelligence with Big Data 출시
  - (Syncsort) 빅데이터 환경에서의 실시간 메인프레임 보안 기능을 제공하는 Ironstream 출시

## &lt;국외 주요 사업자 서비스 동향&gt;

사업자	주요 현황
IBM	- 2012년 12월, 인포스피어 가디엄 포 하둡 발표
Apache Foundation	- 2017년 8월 현재 Atlas라는 하둡을 위한 자동화된 메타데이터 솔루션 (버전 0.9)를 배포
Palantir Technologies	미국 빅데이터 분석 전문 기업으로 2004년 시작하여, 20조 매출을 기록하고 있으며, Palantir Gotham은 데이터를 통합하여 지식관리 기능을 제공하며, Palantir Metropolis는 데이터 통합, 정보관리 정량적 분석을 제공함

## 2.5. IPR 현황 및 전망

### ○ 특허분석 개요

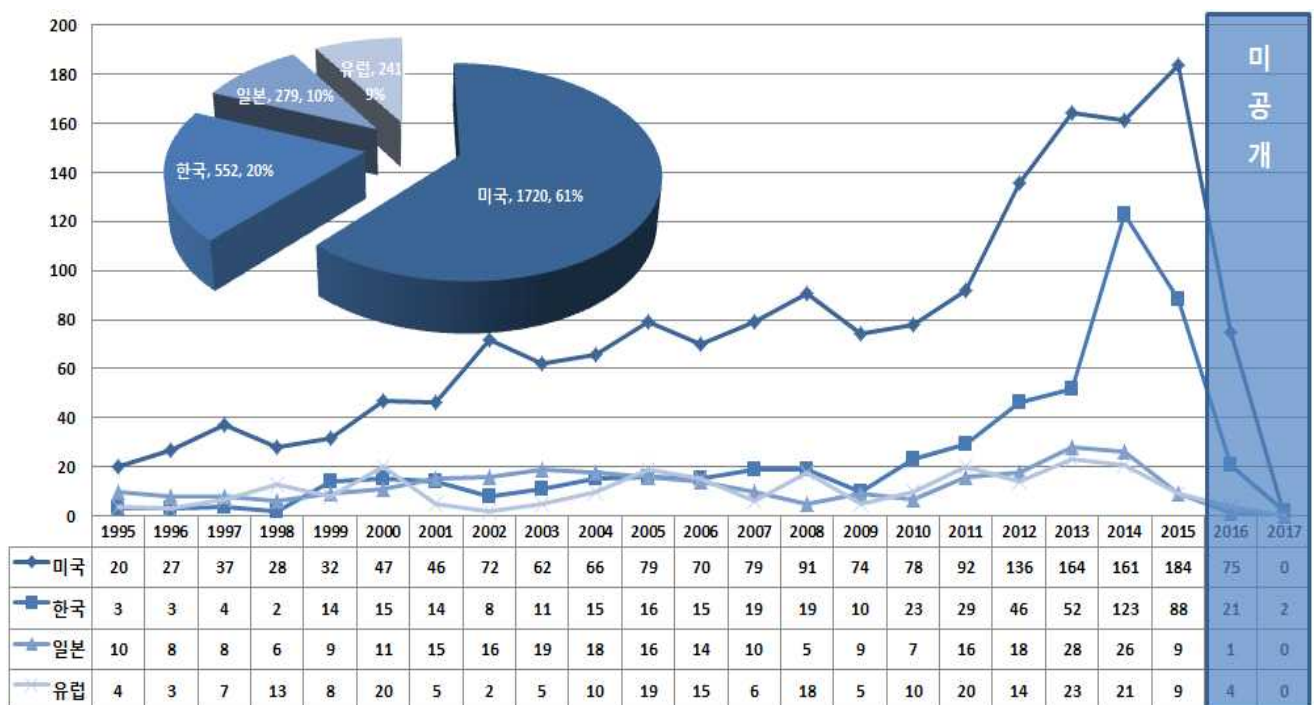
- 빅데이터 분야에 있어서, 2017년 6월 현재까지 한국, 미국, 일본, 유럽에 공개된 특허들을 대상으로 빅데이터 기술과 관련된 총 2,792건의 특허를 대상으로 분석을 수행함

### ○ 연도별 출원 동향

연도	'95	'96	'97	'98	'99	'00	'01	'02	'03	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	합계
출원건수	37	41	56	49	63	93	80	98	97	109	130	114	114	133	98	118	157	214	267	331	290	101	2	2792

- 빅데이터 기술과 관련된 출원은 2004년부터 점차적으로 증가하기 시작하였고, 2013년을 기점으로 급증하고 있는 것으로 조사되었으며, 미공개구간인 2016년도 100건 이상이 특허가 출원되고 있어 급격한 감소 없이 꾸준한 특허출원이 이루어질 것으로 예상됨

### ○ 특허 출원년도별 특허공개 동향



- 빅데이터 기술과 관련하여 출원국가 DB(한국, 미국, 일본, 유럽)별로 공개 및 등록된 특허는 미국특허가 1,720건(61%)으로 가장 많은 특허가 출원되고 있으며, 그 다음으로 한국특허 552건(20%), 일본특허 279건(10%), 유럽특허 241건(9%) 순으로 특허출원이 이루어지고 있음
- 미국의 경우 1990년대 중반부터 특허출원이 이루어져 꾸준한 특허출원량을 보이고 있으며,

2012년들어 특허출원량이 증가하는 양상을 보이고 있음, 한국 역시 2012년들어 출원량이 증가하기 시작하였고 2014년 급격한 출원 증가 양상을 보임, 미공개구간이 2016년의 특허출원량을 고려 시 미국과 한국에서는 급격한 증감 없이 일정량이 특허가 계속 출원될 것으로 예측됨

○ 한국특허에서의 주요 출원인별 출원 현황

순위	출원인	출원건수
1	ETRI	37
2	KISTI	23
3	KT	18
4	SAMSUNG	17
5	SK Telecom	9
6	IBM	8
7	고려대학교	7
8	KAIST	7
9	경희대학교	6
10	서울대학교	6
11	기타	414
합계		552

- 한국 특허청으로는 ETRI가 가장 많은 출원량을 나타내고 있으며, 그 뒤를 KISTI, KT, SAMSUNG, SK Telecom 등의 다수의 특허출원을 수행한 것으로 조사 되었고, 기업과 대학교 위주의 출원이 이루어지고 있음
- 해외 기업으로는 IBM이 유일하고 국내에 8건의 특허출원을 진행하였음

## ○ 해외특허에서의 주요 출원인별 출원 현황

순위	출원인	출원건수
1	IBM	187
2	HITACHI	137
3	MICROSOFT	66
4	NEC	40
5	FUJITSU	38
6	SIEMENS	33
7	SPLUNK	33
8	EMC	29
9	NTT	29
10	GE	23
11	HEWLETT-PACKARD	23
12	TOSHIBA	23
13	GOOGLE	20
14	기타	1559
합계		2240

- 미국, 유럽, 일본 등 해외의 주요 시장에서는 IBM, HITACHI, Microsoft, NEC, FUJITSU 등의 기업이 가장 많은 출원을 하고 있으며, 특히 IBM과 HITACHI는 100건 이상의 특허출원을 진행한 것으로 나타났음
- 가장 많은 특허를 출원 중인 IBM은 국내에서도 다수의 특허를 출원 중에 있으며, HITACHI, NEC, FUJITSU 등 일본 기업이 빅데이터 관련 기술에 많은 투자를 진행하고 있는 것으로 조사되었음

## 2.6. 표준화 현황 및 전망

표준화 수준	국내	<input type="checkbox"/> 기획→ <input type="checkbox"/> 항목승인→ <input type="checkbox"/> 개발/검토→ <input checked="" type="checkbox"/> 최종검토→ <input type="checkbox"/> 제/개정	표준화 격차/특성	0년
	국제	<input type="checkbox"/> 기획→ <input type="checkbox"/> 항목승인→ <input checked="" type="checkbox"/> 개발/검토→ <input type="checkbox"/> 최종검토→ <input type="checkbox"/> 제/개정		후행
* 표준화 특성: 선행(선표준화 후기술개발) - 병행(표준화 & 기술개발 동시추진) - 후행(선택기술개발 후표준화)				

구분	표준화 기구		표준화 현황
국제 (공적)	ISO	TC 184/SC4/WG13	데이터 품질관리 평가 프로세스 참조 모델 표준, 데이터 품질관리 평가 프레임워크 표준, 데이터 품질관리 성숙 모델 표준 개발 중
		WG 9	빅데이터 용어 및 참조구조 표준 개발 중
	JTC 1	SC 27	빅데이터 참조구조의 보안 및 프라이버시 패브릭 표준 개발 중
		SC 32	데이터베이스 언어, 멀티미디어 데이터, 메타데이터 등 데이터 관리 및 교환 관련 표준 개발 중
	ITU-T	SG 13	클라우드 기반 빅데이터 요구사항 및 구조 표준, 빅데이터 프로비던스, 빅데이터 통합, 빅데이터 교환 등의 표준 개발 중
		SG 17	빅데이터 분석을 위한 보안 요구사항 및 프레임워크, 빅데이터로서의 서비스의 보안 가이드라인 개발 모바일 인터넷 환경에서의 빅데이터 보안 요구사항 및 프레임워크, 서비스로서의 빅데이터의 보안 가이드라인 개발
		SG 20	IoT에서 빅데이터를 이용하기 위한 요구사항 표준 개발 중
국제 (사실상)	W3C BigData Community Group		BigData Community Group에서 빅데이터 처리를 위한 표준 구조, 프로그램 API 정의, 상호환성, 보안, 저비용을 보장할 프로그램 언어 개발 등
	ODCA (Open Data Center Alliance)		안전하게 데이터를 수집, 관리, 분석할 수 있는 표준과 기존 BI 솔루션 사이의 상호 운영성을 보장할 수 있는 표준 개발을 추진
	OASIS		빅데이터 환경에서 메시지 프로토콜 및 데이터 프로토콜 등 상호 교환 프로토콜에 관한 정의 및 표준
	OGF		Web Services Data Access and Integration : 데이터 서비스 또는 접근이 가능한 자원을 기술하기 위한 표준으로 데이터 자원에 대한 접근을 지원하기 위한 메시지 패턴을 정의하고 있으며, 웹서비스 기반이지만 데이터 접근과 통합의 필요성 강조
	DMTF		- 1992년 설립하여 Cisco, Dell, EMC, 후지쯔, 히타치, HP, IBM, Intel, MS 등이 참여 중
국내	TTA	빅데이터 SPG	클라우드 기반 빅데이터 서비스를 위한 기능적 요구사항 등 국내 표준 15건의 표준을 제정하여 국내표준화 추진
		메타데이터 PG(PG606)	(메타데이터 표준화) 관계형 지식 추출 모델 및 데이터 포맷 표준화 · TTA.KO-11.0010/R1, 형태소 태깅 말뭉치 작성용 품사 태그세트 · TTA.KO-10.0852, 개체명 태그세트 및 태깅 말뭉치 · TTA.KO-10.0853, 의존 구문분석 말뭉치 구축을 위한 의존관계 태그세트 및 의존관계 설정 방법

### 2.6.1. 국내 표준화 현황 및 전망

○ (빅데이터 개념 및 정의) 빅데이터의 특징 및 기본 개념과 함께, 데이터 제공자, 빅데이터 서비스 제공자, 빅데이터 서비스 사용자로 구성된 빅데이터 생태계 모델을 포함하는 표준의 개발이 TTA SPG22 및 빅데이터 포럼을 통해 2015년 개발 완료

- (TTA 빅데이터 PG(SPG22)) 2014년부터 빅데이터의 전반적 개념 및 생태계 모델을 포함하는 표준 개발이 추진되었으며, 2015년에 표준이 제정되었음

#### < 국내 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA SPG22	TTAK.KO-10.0899-Part1, 빅데이터 프레임워크 - 제1부: 개요 및 정의	2015	빅데이터 개념 및 정의

○ (빅데이터 참조 구조) 빅데이터 생태계 모델을 기반으로 빅데이터 서비스 구현을 위한 기능 요소들을 정의하는 빅데이터 참조 구조에 대한 표준을 TTA 빅데이터 PG를 통해 2016년 개발 완료함

- (TTA 빅데이터 PG(SPG22)) 2014년 착수하여 2016년 표준 제정

#### < 국내 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA SPG22	TTAK.KO-10.0973, 빅데이터 프레임워크 - 제2부: 참조 아키텍처	2016	빅데이터 참조 구조
TTA SPG22	TTAK.KO-10.0705, 클라우드 기반 빅데이터 서비스를 위한 프레임워크 정의	2013	
TTA SPG22	TTAK.KO-10.0706, 클라우드 기반 빅데이터 서비스 참조 모델	2013	

○ (클라우드 기반 빅데이터 서비스 기능 및 아키텍처) 국내에서는 국책연구소(ETRI)를 중심으로 클라우드 기반의 빅데이터 서비스 표준화 연구를 하고 있으며, 국제 표준화 기구(ITU-T) 2015년 11월에 제정된 ITU-T Y.3600에 주 에디터로 참여한데 이어, Q18/13에서 개발하고 있는 Y.BDaaS-arch 구조 표준 개발에 참여 중

- (TTA 빅데이터 SPG) 2015년 11월에 ITU-T에서 Y.3600표준을 제정 완료됨에 따라 SPG21에서는 Y.3600을 준용하는 국내 단체표준 개발하여 2016년 12월에 제정완료 하였음. 제정하였으며, 2016년 12월에 제정을 목표로 하고 있음
- (ETRI) 국제표준화 기구(ITU-T)의 Q18/13에서 관련 표준화 작업(Y.BDaaS-arch)에 참여 중

#### < 국제 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA SPG22	TTAE.IT-Y.3600, “빅데이터 - 클라우드 컴퓨팅 기반의 요구사항 및 기능”	2016	클라우드 기반 빅데이터 서비스 기능 및 아키텍처



- (빅데이터 교환을 위한 프레임워크 및 요구사항) 2017년 국내 데이터 거래소들을 연계하여 빅데이터를 통합적으로 유통하기 위한 논의가 시작되고 있는 상황으로 이를 위한 단계적 기술 적용 방안 및 표준화 대상이 기획되고 있는 단계임
- (빅데이터 프로비넌스 요구사항) 빅데이터 유통과 관련, 빅데이터의 품질을 보장하기 위한 정보로서 데이터의 이력/갱신 정보를 관리하기 위한 방안 등이 논의되고 있으나 표준화는 아직 진행되지 않음
- (빅데이터 환경을 위한 메타데이터) 빅데이터의 상호운용성을 지원하기 위한 메타데이터 표준화 요구가 증대되고 있으나 표준화는 아직 진행되지 않음
  - (TTA 메타데이터 PG(PG606)) 2015년부터 빅데이터 수집, 저장, 분석, 관리 관련 메타데이터 표준을 주요 과제로 하고 있으며, 2018년도에는 구체적인 표준화 작업이 진행 예상
  - (TTA 빅데이터 PG(SPG22)) 빅데이터 상호운용 표준화를 위해 다양한 표준을 활발히 진행해 왔으며, 2018년에는 메타데이터 분야에 대한 표준화 작업이 진행될 것으로 예상

< 국내 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA PG606	TTAK.KO-10.0976, 연구데이터 관리 및 공유를 위한 메타데이터	2017	빅데이터 환경을 위한 메타데이터
TTA PG606	TTAK.KO-11.0224, 과학 실험 데이터(HDF5) 관리와 교환을 위한 메타데이터	2016	
TTA SPG22	TTAE.OT-10.0418, 데이터 분석 모델 공유를 위한 마크업 언어 V.4.2.1	2016	
TTA SPG22	TTAE.OT-10.0417, 웹 환경에서 테이블 데이터를 위한 메타데이터 어휘	2016	

- (빅데이터 품질측정 및 관리모델) 데이터 품질관리를 위한 요소를 정의, TTA PG606 메타데이터 표준화위원회를 통해 개발 중에 있음, 최근 빅데이터 이슈로 인해 빅데이터 품질을 정의하기 위한 메타데이터 표준화에 대한 요구 증대. 국가기술표준원 산업데이터전문위원회에서는 산업데이터 전반에 걸친 표준화를 모니터링하며 산업데이터관련 표준을 KS로 제정
  - (TTA 메타데이터 PG(PG606)) 2009년부터 데이터품질 분야 표준 제정

< 국내 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA PG606	TTAK.KO-10.0474, 데이터 품질 관리 프로세스	2010	빅데이터 품질측정 및 관리모델
TTA PG606	TTAK.KO-10.0475, 데이터 품질 특성	2010	
TTA PG606	TTAK.KO-10.0473, 데이터 관리 프로세스	2010	



## ○ (빅데이터 프라이버시 및 보안관리) 빅데이터 환경에서의 신뢰이슈 표준안 개발 중

## &lt; 국내 표준화 현황 &gt;

개발기구	표준(안)명	개발연도	관련 표준화항목
TTA SPG22	2016-1878, 빅 데이터 환경에서의 신뢰 응용 유즈 케이스 (기술보고서)	개발중 (2018)	빅데이터 프라이버시 및 보안관리
	2016-1876, 빅 데이터 환경에서의 신뢰 이슈 (기술 보고서)	개발중 (2018)	

## 2.6.2. 국외 표준화 현황 및 전망

○ (빅데이터 개념 및 정의) ITU-T는 2015년 빅데이터에 대한 개념 및 생태계 모델에 대한 내용을 포함하는 Y.3600 'Requirements and capabilities for cloud computing based big data'를 공적 표준화기구 최초로 개발 완료하였으며, ISO/IEC JTC 1에서는 2015년 3월부터 WG 9 Big Data를 통해 빅데이터 개념 및 어휘에 대한 표준화를 진행

- (ITU-T SG13) 빅데이터 생태계 모델을 통해 이에 참여하는 역할 및 활동에 대한 정의가 2015년 개발 완료
- (ISO/IEC JTC 1 WG 9) 2014년부터 빅데이터 개요 및 용어에 대한 표준 개발이 추진 중에 있으며, 2018년까지 개발 완료를 목표로 2017년 7월 현재 CD3 (3rd Committee Draft) 투표가 완료

## &lt; 국제 표준화 현황 &gt;

개발기구	표준(안)명	개발연도	관련 표준화항목
JTC 1 WG 9	ISO/IEC 20546, Information technology - Big data - Overview and vocabulary	진행 중 (2018)	빅데이터 개념 및 정의
ITU-T SG13	ITU-T Y.3600, Requirements and capabilities for cloud computing based big data	2015	

○ (빅데이터 참조 구조) 빅데이터 생태계 모델을 기반으로 빅데이터 서비스 구현을 위한 기능 요소들을 정의하는 빅데이터 참조 구조에 대한 표준을 TTA 빅데이터 PG를 통해 2016년 개발 완료

- (ISO/IEC JTC 1 WG 9) 2014년에 ISO/IEC 20547 빅데이터 - 참조 아키텍처 개발을 시작하였으며, 2016년에는 다섯 개의 개별 문서로 분리(개요 및 적용 절차, 유즈케이스 및 요구사항, 참조 아키텍처, 보안 요구사항, 표준 로드맵)되어 각각 개발을 진행 중
- (ITU-T SG13) 2017년 2월 회의를 통해 개발 중인 클라우드 기반의 빅데이터 아키텍처를 빅데이터 참조 아키텍처와 클라우드 기반 아키텍처의 두 가지 문서로 분리하여 개발을 진행 할 것이 논의됨

## &lt; 국제 표준화 현황 &gt;

개발기구	표준(안)명	개발연도	관련 표준화항목
JTC 1 WG 9	ISO/IEC 20547-1, Information technology - Big data - Reference architecture - Overview and application process	진행 중 (2018)	빅데이터 참조 구조
	ISO/IEC 20547-2, Information technology - Big data - Reference architecture - Use cases and derived technical considerations	진행 중 (2017)	
	ISO/IEC 20547-3, Information technology - Big data - Reference architecture - Reference architecture	진행 중 (2019)	
	ISO/IEC 20547-4, Information technology - Big data - Reference architecture - Privacy and security fabric	진행 중 (2018)	
	ISO/IEC 20547-5, Information technology - Big data - Reference architecture - Standards roadmap	진행 중 (2017)	

○ (클라우드 기반 빅데이터 서비스 기능 및 아키텍처) ITU-T SG13에서 클라우드 기반 빅데이터의 요구사항을 제정(Y.3600, Big data - Cloud computing based requirements and capabilities)에 이어 2015년 말에 관련 구조 표준(Y.BDaaS-arch)에 대한 작업을 시작하였음. 현재 추진 일정상 2018년도 중반 이후에 표준 제정이 완료될 것으로 예상

- (ITU-T SG13 Q17) 빅데이터 기반 서비스 기능 및 아키텍처의 기반이 되는 Y.3600(Big data - Cloud computing based requirements and capabilities) 표준을 제정 완료(2015년 11월)
- (ITU-T SG13 Q18) 빅데이터 기능 및 구조 표준(Y.BDaaS-arch, “Big data - Functional architecture of Big Data as a Service”) 표준 권고안 개발 작업을 진행중에 있음. 2018년 표준 제정을 완료를 예정

## &lt; 국제 표준화 현황 &gt;

개발기구	표준(안)명	개발연도	관련 표준화항목
ITU-T SG13	Y.3600, “Big data - Cloud computing based requirements and capabilities”	2015	클라우드 기반 빅데이터 서비스 기능 및 아키텍처
	Y.BDaaS-arch, “Big data - Functional architecture of Big Data as a Service”	진행 중 (2018)	

○ (빅데이터 교환을 위한 프레임워크 및 요구사항) 공공 데이터 개방과 함께 누구든지 자유롭게 데이터에 접근하여 활용하는 것을 지원하기 위한 데이터 발간, 데이터 검색, 데이터 유통 및 판매 등을 위한 요구사항 표준이 개발 중

- (ITU-T SG 13) ITU-T SG 13은 2015년 4월부터 이통사들을 중심으로한 빅데이터 유통을 위한 및 프레임워크 및 기능적 요구사항 개발을 추진 중

## &lt; 국제 표준화 현황 &gt;

개발기구	표준(안)명	개발연도	관련 표준화항목
ITU-T SG13	Y.BigdataEX-reqts, Big data exchange framework and requirements	진행 중 (2017)	빅데이터 교환을 위한 프레임워크 및 요구사항

○ (빅데이터 프로비넌스 요구사항) 빅데이터 교환 시, 데이터 출처 및 데이터 처리에 사용된 기능 정보 등을 추적하여 그 이력을 관리하기 위한 표준화가 ITU-T SG13을 통해 진행 중

- (ITU-T SG 13) 2016년 6월 회의를 통해 2018년 표준 제정을 목표로 '빅데이터 프로비넌스 요구사항 (Y.bdp-reqts, Big data - Requirements for data provenance)' 표준 개발을 시작함

< 국제 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
ITU-T SG13	Y.bdp-reqts, Big data - Requirements for data provenance	진행 중 (2018)	빅데이터 프로비넌스 요구사항

○ (빅데이터 환경을 위한 메타데이터) 메타데이터에 대한 표준화 필요성에 대해 공감하고, 이를 위한 표준화 논의와 기초적인 표준 개발 중

- (ITU-T SG13) 2017년 7월 회의에 “빅데이터-메타데이터 프레임워크 및 개념 모델” 표준을 제안하여 빅데이터를 위한 메타데이터 표준을 주도하고 있음

< 국제 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
ITU-T SG13	Y.bdm-sch, Big Data - Metadata framework and conceptual model	진행 중 (2019)	빅데이터 환경을 위한 메타데이터

○ (빅데이터 품질측정 및 관리모델) 빅데이터 활용에 있어서 품질이 중요한 요인이 됨에 따라, 빅데이터의 품질을 측정하기 위한 빅데이터 품질요소, 품질측정 방안 등에 대한 표준 개발 중

- (ISO/IEC JTC1/SC32/WG2) SC32/WG2 메타데이터 작업반에서는 빅데이터 품질 AHG(Ad-Hoc Group)를 만들어 빅데이터의 품질을 정의하는 메타데이터 요소에 대한 기본 개념 표준을 개발 중
- (ISO TC 184/SC4/WG13) WG13은 데이터 품질 표준(ISO 8000)을 만들고 있으며 최근에는 데이터 품질 관리 프로세스 평가 모델 표준을 제정하고 있음

< 국제 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
JTC1/SC32/WG2	ISO/IEC 19583-1, Information technology - Concepts and usage of metadata - Metadata concepts	진행 중 (2018)	빅데이터 품질측정 및 관리모델
ISO TC 184/SC4/WG13	ISO 8000-66, 스마트 공장을 위한 데이터 품질 관리 프로세스 평가 모델	진행 중 (2019)	
	ISO 8000-61, 데이터 품질관리 프로세스 참조 모델	2016	

- (빅데이터 프라이버시 및 보안관리) 빅데이터 보안 분야는 ISO에서 빅데이터 참조 구조의 여러 계층에서 수행해야 하는 프라이버시 및 보안 측면을 위한 표준인 ISO/IEC 27047-4를 개발 중이고 ITU-T에서는 서비스로서의 빅데이터에 필요한 보안 요구사항 명세인 X.GSBDaaS 를 개발 중

< 국제 표준화 현황 >

개발기구	표준(안)명	개발연도	관련 표준화항목
ISO/IEC	ISO/IEC 20547-4, Information technology - Big Data Reference Architecture - Part 4: Security and Privacy Fabric	진행 중 (2019)	빅데이터 프라이버시 및 보안관리
ITU-T	ITU-T X.GSBDaaS, Guidelines on security of Big Data as a Service	진행 중 (2019)	
	ITU-T X.srfb, Security requirements and framework for Big Data Analytics in mobile Internet services	진행 중 (2019)	

## 2.7. 오픈소스 현황 및 전망

### ○ 하둡 (Hadoop)

- 대용량의 데이터를 저장하고 처리할 수 있는 분산 시스템으로 HDFS, 맵리듀스, YARN으로 구성됨
- 하둡-HDFS : HDFS는 데이터를 다수의 노드에 분산해서 저장하여 스케일아웃에 용이하고 전통적인 스토리지의 저장 용량의 한계를 넘을 수 있음. HDFS는 대용량 파일의 스트리밍 읽기와 쓰기에 뛰어난 성능을 보임
- 하둡-맵리듀스 : 맵리듀스는 프로그래밍 모델과 프레임워크 구현 환경으로 구성. 맵리듀스는 대용량, 분산, 내고장성 데이터 처리 프로그램을 개발자가 쉽게 개발하도록 설계한 최적의 애플리케이션 기술임. 개발자가 맵리듀스에서 맵 함수와 리듀스 함수를 조합하여 코드를 작성하면, 프레임워크는 작업의 병렬처리, 워커 머신의 태스크 스케줄링, 모니터링, 장애 복구 등의 세부사항을 알아서 처리됨
- 하둡-YARN : 리소스 관리 플랫폼인 YARN은 하둡 클러스터의 각 어플리케이션에 필요한 리소스를 할당하고 모니터링하는 업무에 집중함으로써 다양한 어플리케이션이 하둡 클러스터의 리소스를 공유할 수 있도록 하는 핵심 요소임

### ○ 하둡 에코시스템 - 수집 기술

- 하둡 기반의 데이터 수집 기술로 대용량의 데이터를 저장하고 처리할 수 있는 소프트웨어로는 플룸, 스쿱, 카프카 등이 있음
- 플룸 : 플룸은 에이전트가 설치된 머신에서 데이터 소스를 수집하여 하둡 분산 환경의 콜렉터에 전송하고 이를 로컬파일시스템, 하둡분산파일시스템, 하이브, HBase 등에 저장하는 데이터 수집 엔진임. 다수의 웹 서버에서 로그파일을 수집하고 해당 파일의 로그 이벤트를 처리하기 위해 HDFS에 위치한 새로운 통합 파일로 옮기는 것은 플룸을 사용하는 적형적인 예임. 일반적인 최종 목적지는 HDFS지만 플룸은 HBase나 솔라와 같은 다른 시스템에도 이벤트를 기록할 수 있음
- 스쿱 : 스쿱은 구조화된 데이터 저장소에서 데이터를 추출해서 하둡으로 보내 처리할 수 있도록 해주는 오픈 소스 도구임. 이러한 처리는 맵리듀스 프로그램이나 하이브 같은 다른 고차원 도구로도 할 수 있음. 또한 관계형 데이터베이스의 데이터를 HBASE로 옮기는 데 스쿱을 사용할 수 있다. 분석 파이프라인을 거쳐 최종 산출물이 나오면 스쿱은 이 산출물을 다시 데이터 저장소로 보내 다른 사용자가 사용할 수 있도록 해줌
- 카프카 : 카프카는 확장성이 좋고 처리량이 높은 분산 메시지 시스템으로 우수한 메시지 전달 성능을 보장함. 링크드인의 실시간 데이터나 오프라인 데이터를 처리하고 수집하는데 사용됨. 카프카는 높은 처리량을 유지하면서 병렬처리에 초점을 맞춘 구조이며 다량의 클라이언트 생성이 가능한 유연성을 제공하고 분산처리를 지원함

## ○ 하둡 에코시스템 - 저장 및 처리 기술

- 하둡 기반의 데이터 저장 및 처리 기술로는 피그, 하이브가 있음
  - 피그 : 맵리듀스 프로그래밍은 저수준 언어인데 비해 피그는 고수준 언어임. 피그 라틴이라는 스크립트 언어로 코드를 작성하면 피그 엔진은 이를 맵리듀스 프로그램으로 변환하고 하둡 상에서 실행하고 그 결과를 알려줌
  - 하이브 : 하이브는 하둡 기반의 데이터 웨어하우징 프레임 워크로, 빠른 속도로 성장하는 페이스북의 소셜 네트워크에서 매일같이 생산되는 대량의 데이터를 관리하고 학습하기 위해 개발됨. 하이브는 데이터의 스키마를 관리할 수 있으며, 스크립트 언어가 아닌 SQL 언어를 지원하므로 기존 데이터베이스 프로그래머들이 쉽게 하이브 쿼리를 익혀서 빨리 개발할 수 있어, 하이브는 하둡 기반의 데이터웨어하우스로 불림
- HDFS에 저장된 데이터를 SQL 혹은 SQL과 유사한 형태로 처리를 요청하고 분산 처리하는 시스템을 통칭하는 것이 SQL On Hadoop임. 국내에서 개발한 타조와 클라우데라의 임팔라 등 다양한 도구가 있음
  - Hive on Tez : 호튼웍스의 스팅거 프로젝트로 하이브를 발전시키고 성능을 강화한 분석 엔진이다. 새로운 맵리듀스인 테즈를 기반으로 개발됨
  - 드릴(Drill) : 구글의 Dremel(일명 BigQuery)의 오픈소스 버전으로 높은 확장성과 하둡/NoSQL과의 쿼리 인터페이스 지원이 특징임
  - SparkSQL(Shark) : 스파크(Spark)는 인메모리 기반의 초고속 분산병렬처리 프레임워크로 SparkSQL은 하이브에 저장된 데이터에 SQL 질의를 할 수 있는 것이 특징임
  - 임팔라(Impala) : 하이브 사용자를 고려한 빠른 SQL 쿼리 엔진이다. HDFS나 HBase에 저장된 데이터를 대상으로 분석, 문법은 HiveQL과 동일
  - 프레스토(Presto) : 페이스북에서 개발했으며, 빠른 속도의 검색과 집계 가능

## ○ 하둡 에코시스템 - 분석 기술

- 하둡 기반의 분석 기술로는 머하웃과 스파크가 있음
  - 스파크 : 스파크는 대용량 데이터 처리를 위한 클러스터 컴퓨팅 프레임워크다. 스파크는 실행 엔진으로 맵리듀스를 사용하지 않는 대신 스파크는 클러스터 기반으로 작업을 실행하는 자체 분산 런타임 엔진이 있음. 하지만 스파크는 이 장에서 살펴볼 API와 수행 측면 모두에서 맵리듀스와 매우 비슷함. 스파크는 하둡과 밀접하게 통합되어 있어서 YARN기반으로 실행할 수 있고, 하둡 파일 포맷과 HDFS 같은 기반 저장소를 지원함
  - 머하웃(Mahout) : 머하웃은 추천시스템, 기계 학습 등 데이터 마이닝 알고리즘을 쉽게 구현하도록 지원하는 솔루션으로 하둡을 기반으로 구현되어 있음

## ○ 하둡 클러스터 관리 및 운영 기술

- 하둡 클러스터를 관리하고 운영하는 기술로는 주키퍼, 우지 등이 있음
  - 주키퍼(Zookeeper) : 주키퍼는 분산 환경에서 다수의 머신과 서비스간의 상호 조정 역할을 함. HBase의 필수 구성요소이며, 하둡 2.0 버전의 네임노드 이중화에도 사용됨. 주키퍼의 핵심은 분산 데이터 저장소임. 다수의 머신에 동일한 데이터를 저장하며 고가용성과 빠른 응답속도를 보장함. 주키퍼의 데이터 동기화와 클라이언트와의 세션관리 및 이벤트 감시자 기능을 활용하면 분산 시스템에서 다양한 머신과 서비스의 상호조정이 가능함
  - 에이브로 : 에이브로는 특정 언어에 종속되지 않는 언어 중립적 데이터 직렬화 시스템이다. 하둡의 창시자인 더그 커팅이 하둡 Writable(직렬화 방식)의 주요 단점인 언어 이식성을 해결하기 위해 만든 프로젝트다. 에이브로는 현재 C, C++, C#, 자바, 자바스크립트, 펄, PHP, 파이썬, 루비 등의 언어로 처리할 수 있는 새로운 데이터 포맷을 제공하기 때문에 단일 언어에 얽매이지 않고 다양한 언어로 데이터셋을 쉽게 공유할 수 있음
  - 우지(Oozie) : 우지는 하둡 맵리듀스 잡, 스쿱 임포트 및 익스포트, 하이브 잡, 피그 잡 등 하둡과 하둡 에코시스템의 작업을 관리하는 워크 플로우 엔진이자 코디네이터 시스템이다. 다양한 하둡 에코시스템의 작업을 주기적으로 실행하고 그 결과를 통보해줌
  - 크런치 : 크런치는 맵리듀스 파이프라인을 작성하는 고수준 API이다. 맵리듀스 대신 크런치를 사용하면 String이나 POJO(plain old Java object)와 같은 프로그래머에게 익숙한 자바 자료형, 풍부한 데이터 변환 기능, 여러 단계의 파이프라인에 집중할 수 있다는 장점이 있음

## ○ R - 분석 기술

- R은 프로그래밍 언어이자 통계적 컴퓨팅을 지원하는 소프트웨어 환경이다. 빅데이터의 등장과 함께 R은 데이터들을 통계적으로 분석하기 위한 도구로 활용되고 있음
  - R의 상용 버전: R은 레드햇이 리눅스를 지원하는 방식과 유사한 서비스와 지원 모델을 추구하는 레볼루션 애널리틱스(Revolution Analytics)로 상용 버전이 배포되고 있음
  - R의 공개 버전: GNU 제너럴 퍼블릭 라이선스(GNU General Public license)에서 공개 버전은 오픈 소스로 쉽게 사용이 가능함
  - R의 진화: R언어가 빅데이터 분야에 활용되면서 단순 통계계산 영역에서 벗어나, 머신러닝과 딥러닝을 포함한 기능도 제공하고 있다. 또한, 빅데이터의 활용을 지원하기 위한 시각화 기술도 제공하고 있음



### Ⅲ. 국내외 표준화 추진전략

#### 3.1. 표준화 SWOT 분석

국내역량요인		강점요인 (S)		약점요인 (W)	
		시장	<ul style="list-style-type: none"><li>- IoT, 모바일 등의 다양한 기기를 통한 데이터 생성 증가</li><li>- 네트워크 및 단말 환경이 최고 수준이라 이에 따른 데이터 생성이 용이</li></ul>	시장	<ul style="list-style-type: none"><li>- 데이터 자체에 대해서 비용을 지불하지 않고 획득하려는 인식</li></ul>
		기술	<ul style="list-style-type: none"><li>- 정부의 지속적인 빅데이터 관련 기술 개발 투자</li></ul>	기술	<ul style="list-style-type: none"><li>- 개인정보보호에 대한 강력한 규제에 인하여 사용 가치가 있는 데이터 확보의 어려움</li></ul>
		표준	<ul style="list-style-type: none"><li>- 국제 표준화 주도 (의장, 에디터 등의 의장단 확보)</li></ul>	표준	<ul style="list-style-type: none"><li>- 국내 기업의 표준화 미참여</li></ul>
국외환경요인					
기회요인 (O)	시장	<ul style="list-style-type: none"><li>- 클라우드, 인공지능 등의 관련 기술과 연계한 빅데이터 활용 가치 증대</li></ul>	<b>【SO전략】</b>  -(시장)데이터 증가 및 데이터 활용 증가로 인한 데이터 확보 용이 -(기술) 오픈 소스 기반의 빅데이터 관련 기술 확보 용이 -(표준) 국제 표준화 의장단 확보 및 표준화 기반 조성		<b>【WO전략】</b>  -(시장) 활용할 데이터는 다수 확보할 수 있으나 고품질의 데이터 확보는 애로사항이 있음 -(기술)빅데이터 기반 기술 개발에 주력하고 있으며, 데이터 확보에 대한 기술 개발이 상대적으로 미약 -(표준)표준 참여를 위한 더 많은 인력이 요구되고 있으며, 이를 통한 표준 경쟁력 강화
	기술	<ul style="list-style-type: none"><li>- 다양한 빅데이터 관련 오픈소스로 기술 장벽이 낮음</li></ul>			
	표준	<ul style="list-style-type: none"><li>- 공적 표준화 기구, 사실 표준화 기구 등에서 빅데이터 표준 개발 중</li></ul>			
위협요인 (T)	시장	<ul style="list-style-type: none"><li>- 특정 솔루션이 지배적이라 솔루션에 대한 시장적 Lock-in 현상 발생 우려</li></ul>	<b>【ST전략】</b>  -(시장)빅데이터 유통을 위한 규제 완화 필요 -(기술)하둠에 의존적이지 않도록 빅데이터 개발에 대한 다양한 생태계 구축 필요 -(표준)데이터 유통 관련 표준 개발 필요		<b>【WT전략】</b>  -(시장) 관련 기술과 연계한 융합 산업으로 빅데이터의 시장 확산 노력 필요 -(기술)개인정보에 대한 안전한 빅데이터 기술 개발 필요 -(표준) 표준전문가 육성 및 핵심 표준기술 보유기관과의 전략적 공조로 지속적인 국제표준 확보
	기술	<ul style="list-style-type: none"><li>- 특정 지배적인 솔루션에 의해 특정 솔루션에 대한 기술적 Lock-in 현상 발생 우려</li></ul>			
	표준	<ul style="list-style-type: none"><li>- 1-2위 글로벌 벤더의 표준화 미참여</li></ul>			
표준화 추진상의 문제점 및 현안 사항					
<ul style="list-style-type: none"><li>- 공적 표준화 기구와 사실 표준화 기구에서 데이터 관련 표준화는 추진되고 있으나, 국내외적으로 이에 대한 참여는 미미한편</li><li>- 특히, 국내의 참여는 매우 저조하며, 연구계와 학계 일부가 표준 개발을 주도하고 있으며, 산업계는 표준보다는 기술 개발 및 서비스 개발에 중점을 두고 있음</li><li>- 빅데이터가 다양한 산업군에서 활용되기 때문에 각 산업군에서 빅데이터를 활용하는 형태의 표준은 파악이 제대로 되지 않고 있음</li></ul>					



### 3.2. 중점 표준화 항목별 국내외 추진전략

○ 선행(선표준화 후기술개발), ㉠ 병행(표준화&기술개발 병행추진), ● 후행(선기술개발 후표준화)

High	< 차세대공략 항목(신규제안) >		< 적극공략 항목(선도경쟁) >	
	㉠ 빅데이터 환경을 위한 메타데이터 ● 빅데이터 프라이버시 및 보안 관리 ㉠ 빅데이터 교환을 위한 프레임워크 및 요구사항 ㉠ 빅데이터 프로비넌스 요구사항		㉠ 빅데이터 참조 구조 ● 빅데이터 품질측정 및 관리모델 ● 클라우드기반 빅데이터 서비스 기능 아키텍처	
Low	< 전략적수용 항목(수용/적용) >		< 다각화협력 항목(부분협력) >	
	● 빅데이터 개념 및 정의			
Low		국내 역량 (표준화/기술개발 수준, 국제 표준화에 국내 기여도 등)		High

#### ○ 영역별 특징 및 대응전략

- **차세대공략 항목(신규제안)** : 미래 핵심기술 및 유망서비스 관련 선행적 표준화 분야  
: 국제표준 기획단계부터 주도적 참여를 통해 국제표준화 선도기반 확보  
: 기술 및 특허 반영을 위한 원천기술 개발 병행 (기술개발-표준화 연계 강화)
- **적극공략 항목(선도경쟁)** : 아직 국제표준 완성도가 낮아 국제표준 선도경쟁이 치열한 분야  
: 국내 기술의 국제표준 반영을 위한 표준화 활동 강화  
: 전략적 대외협력 강화 및 제후를 통한 기술/표준의 Catch-up 전략 추진
- **다각화협력 항목(부분협력)** : 시장에서의 기술/상용화 경쟁이 치열한 분야로 포럼/컨소시엄 위주의 표준화가 진행되는 분야  
: 세계 사실표준화기구 대응 및 국내 포럼 활동 강화  
: 사실표준화기구와 공식표준화기구에 다각적인 대응 모색
- **전략적수용 항목(수용/적용)** : 기술개발 및 국제표준화가 거의 완료단계이고, 서비스/시장 확산을 위한 후속 표준화가 필요한 분야  
: 국제표준의 수용/적용을 통한 국제 호환성 확보 및 국내 시장 확산  
: 킬러 애플리케이션/서비스 개발과 병행하여 틈새표준 발굴, 표준화 추진

(전략적 수용   후행) 빅데이터 개념 및 정의							
전략적 중요도 / 국내 역량	<p>국제표준화 국내 기여도</p> <p>정책 부합성</p> <p>시장/기술적 파급효과</p> <p>IPR 확보 가능성</p> <p>국외대비 국내 표준화 역량</p> <p>국외대비 국내 기술개발 수준</p> <p>앞선</p> <p>비슷</p> <p>뒤집</p> <p>높음</p> <p>보통</p> <p>낮음</p>				표준화 기구/ 단체	국내	TTA 빅데이터 SPG
	국외	JTC1 WG9, ITU-T SG13					
	국내 참여 업체/ 기관	ETRI					
기술 개발 단계	국내	□기초연구→□실험→■시작품→□제품화→□사업화		기술 수준	90% (선도국가대비)		
	국외	□기초연구→□실험→■시작품→□제품화→□사업화		기술 격차	1.0년		
	선도국가/ 기업	미국 / Google, 아마존, MS, IBM, SAP, SAS, 오라클 한국 / ETRI					
표준화 단계	국내	□과제기획→□과제승인→□개발→□검토→■표준채택		표준 수준	100 (선도국가대비)		
	국제	□과제기획→□과제승인→■개발→□검토→□표준채택		표준 격차	0년		
	선도국가/ 기업	한국/ETRI, 미국/NIST					
<p>- Trace Tracking : 방어적수용(Ver.2017) → 전략적수용(Ver.2018)</p> <p>빅데이터 개념 및 정의는 ISO/IEC JTC1 WG9가 중심이 되어 2018년 제정을 목표로 개발이 진행 중으로, 국내 표준화는 완료되었으며, 국제표준과의 부합화 관점에서 Ver.2017에서도 전략적수용 항목으로 분류</p>							



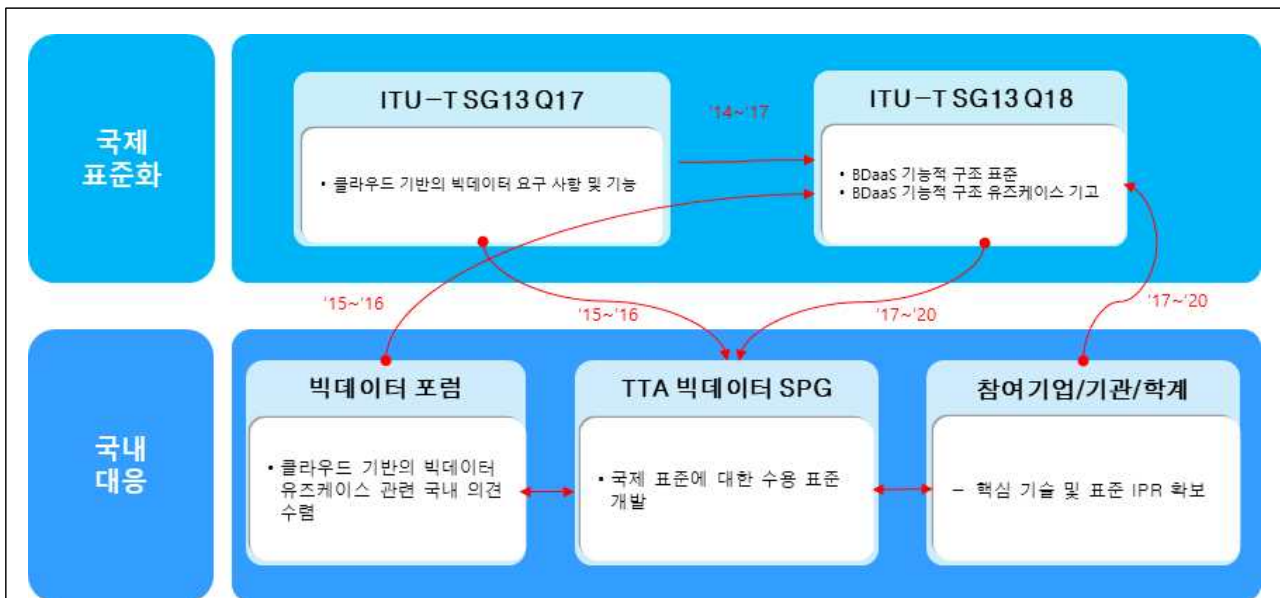
## (적극공략 | 병행) 빅데이터 참조 구조

전략적 중요도 / 국내 역량				표준화 기구/ 단체	국내	TTA 빅데이터 SPG
					국외	JTC1 WG9, ITU-T SG13
					국내 참여 업체/ 기관	ETRI
기술 개발 단계	국내	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input checked="" type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 수준	90% (선도국가대비)	
	국외	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input type="checkbox"/> 시작품→ <input checked="" type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 격차	1.0년	
	선도국가/ 기업	미국 / Google, 아마존, IBM, 오라클 한국 / 다음소프트, SKT				
표준화 단계	국내	<input type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input checked="" type="checkbox"/> 표준채택		표준 수준	100% (선도국가대비)	
	국제	<input type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input checked="" type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 격차	0년	
	선도국가 /기업	한국/ETRI, 미국/NIST				
<p>- Trace Tracking : 적극공략(Ver.2017) → 적극공략(Ver.2018)</p> <p>빅데이터 참조 구조는 ISO/IEC JTC1 WG9가 중심이 되어 2015년부터 개발이 진행되고 있으며, 우리나라가 에디터로 참여함에 따라 우리나라 주도의 개발이 가능하며 시장/기술 파급력이 높은 항목으로 판단되어 Ver.2018에서도 적극공략 항목으로 분류</p>						



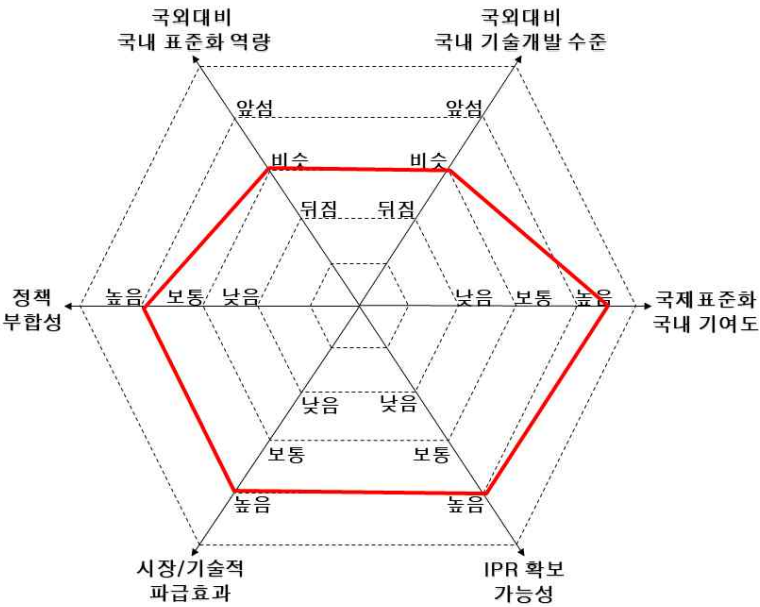
## (적극공략 | 후행) 클라우드 기반의 빅데이터 서비스 기능 및 아키텍처

전략적 중요도 / 국내 역량	<p>외국대비 국내 표준화 역량</p> <p>외국대비 국내 기술개발 수준</p> <p>앞섬</p> <p>비슷</p> <p>뒤짐</p> <p>정책 부합성</p> <p>높음</p> <p>보통</p> <p>낮음</p> <p>국제 표준화 국내 기여도</p> <p>시장/기술적 파급효과</p> <p>IPR 확보 가능성</p> <p>높음</p> <p>보통</p> <p>낮음</p>			표준화 기구/ 단체	국내	TTA 빅데이터 (SPG22), 빅데이터 포럼
	국제	ITU-T SG13				
	국내 참여 업체/ 기관	ETRI				
기술 개발 단계	국내	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input checked="" type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 수준	80% (선도국가대비)	
	국외	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input checked="" type="checkbox"/> 사업화		기술 격차	2년	
	선도국가/ 기업	미국/구글, 아마존, 페이스북 한국/모비젠, 인브레인				
표준화 단계	국내	<input checked="" type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 수준	90% (선도국가대비)	
	국제	<input type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input checked="" type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 격차	1년	
	선도국가/ 기업	한국/ETRI, 중국/차이나텔레콤, 폴란드/Orange polska				
<p>- Trace Tracking : 적극공략(Ver.2017) → 적극공략(Ver.2018)</p> <p>클라우드 기반의 빅데이터 서비스 기능 및 아키텍처 분야는 우리나라의 표준화 기여도 및 시장 및 기술 참여도가 높은 항목으로 2017년도에 이어 2018년에도 적극 공략으로 분류함</p>						

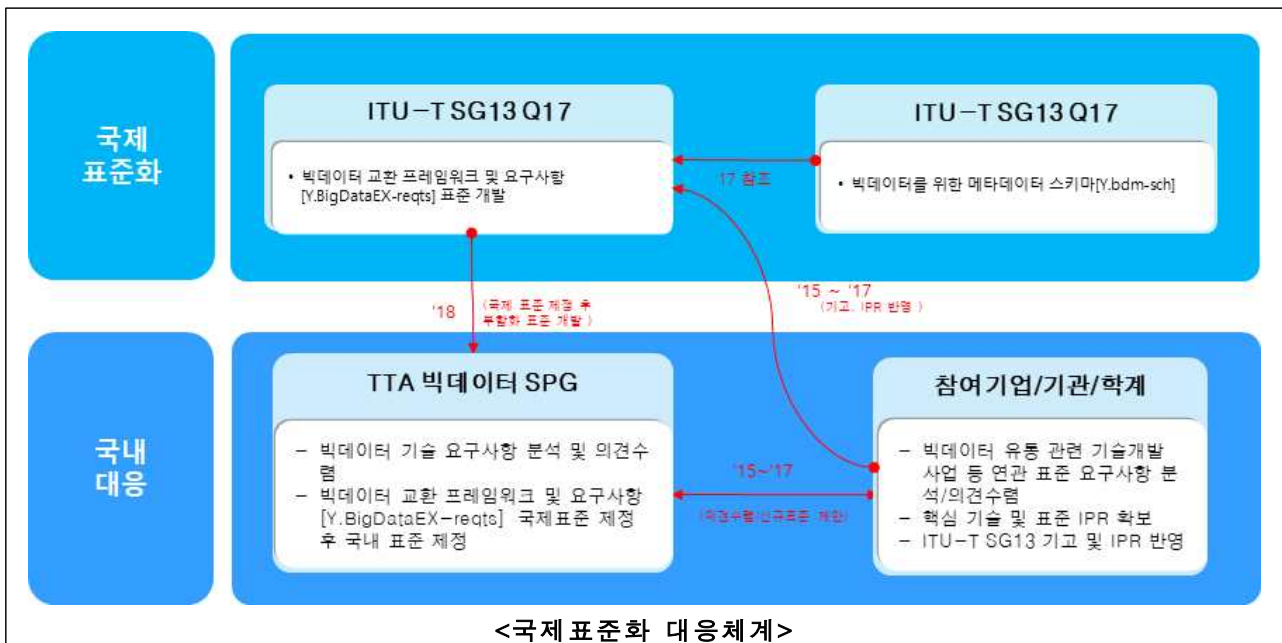


&lt;국제 표준화 대응체계&gt;

국제 표준화 대응 방안	<p><b>&lt;현황&gt;</b></p> <ul style="list-style-type: none"> <li>- ITU-T SG13에서 클라우드 기반의 빅데이터 서비스 기능에 대한 표준 Y.3600(Requirements and capabilities for cloud computing based big data)표준안이 2015년 7월 제정하였으며, 후속으로 ‘클라우드 기반의 빅데이터 기능 및 구조 표준(Y.BDaaS-arch)’을 ITU-T SG13 Q18에서 개발 중에 있음</li> </ul> <p><b>&lt;대응방안&gt;</b></p> <ul style="list-style-type: none"> <li>- (국제 표준화기구 국내 산학연 표준 전문가 참여) 국내에서는 현재 에디터로 참석하고 있는 Y.BDaaS-arch권고안 개발에 지속적인 기고 활동을 통해 국내 관련 포럼 및 TTA 빅데이터 PG(SPG22)에서 제시된 의견을 최대한 반영하고, 추가적인 신규 표준 항목을 도출하여 국제 표준을 지속적으로 추진할 필요가 있음</li> </ul>
국내 표준화 추진 계획	<p><b>&lt;현황&gt;</b></p> <ul style="list-style-type: none"> <li>- 현재 국책기관인 ETRI를 중심으로 ITU-T SG13 Q18에서 국제 표준 개발(클라우드 기반의 빅데이터 기능 및 구조 표준)에 적극 참여하고 있음. 그러나 산업계의 참여는 미진한 상태임</li> </ul> <p><b>&lt;추진계획&gt;</b></p> <ul style="list-style-type: none"> <li>- (학계나 기업과의 협력을 통해 표준에 국내의 기술을 적극 반영) 현재 진행 중인 국제표준 개발에 한국에서 에디터로 참여하여 활발히 활동하고 있음. 이에 국내의 산업체 등의 의견을 수렴하여 국내의 의견을 적극적으로 반영할 예정임</li> <li>- (국내 표준화 추진) ITU-T SG13에서 개발되고 있는 국제 표준화 병행해서 TTA SPG22에서 국내 단체표준으로 제정을 추진함</li> </ul>
표준 특 히 전략	<ul style="list-style-type: none"> <li>- <b>표준 초종기 및 R&amp;D 중후기 전략 : 표준 필수특허 설계전략</b></li> <li>- (세부전략) 시나리오를 중심으로 한 클라우드 기반의 빅데이터 서비스 방법 관련 기술들에 대한 특허 출원과 함께 이를 유즈케이스 형태로 표준의 내용에 반영</li> </ul>
기술개발 -표준화 -IPR 연계 방안	<ul style="list-style-type: none"> <li>- 기술개발 후속 표준화</li> <li>- (세부전략) 현재 개발되어 있는 기술을 중심으로 이미 시작된 국제 표준 개발에 국내의 기술을 반영함</li> </ul>

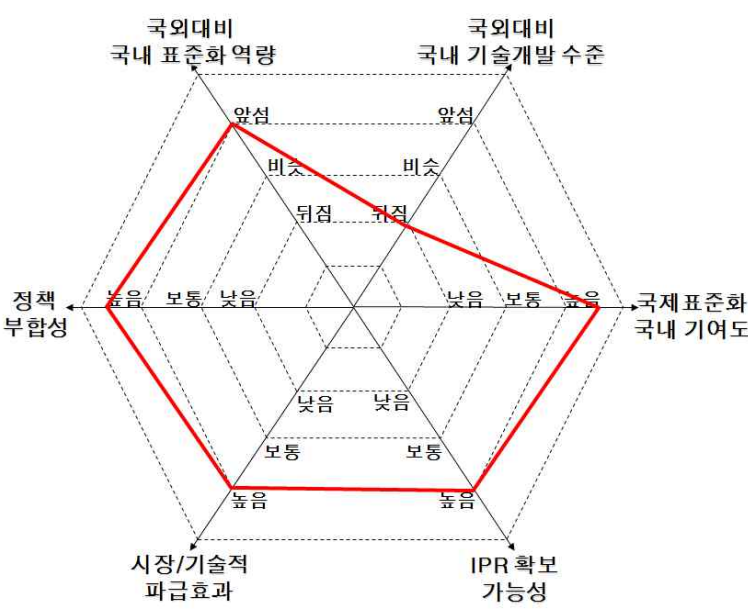
(차세대공약   병행) 빅데이터 교환을 위한 프레임워크 및 요구사항						
전략적 중요도 / 국내 역량				표준 화 기구/ 단체	국내	TTA 빅데이터 SPG
					국 외	ITU-T SG13
					국내 참여 업체/ 기관	ETRI
기술 개발 단계	국내	□기초연구→□실험→■시작품→□제품화→□사업화		기술 수준	90% (선도국가대비)	
	국 외	□기초연구→□실험→■시작품→□제품화→□사업화		기술 격차	1년	
	선도국가/ 기업	미국 / BDEX 한국 / KDB				
표준화 단계	국내	■과제기획→□과제승인→□개발→□검토→□표준채택		표준 수준	90% (선도국가대비)	
	국제	□과제기획→□과제승인→■개발→□검토→□표준채택		표준 격차	1년	
	선도국가 /기업	한국/ETRI, 미국/NIST				
<div>- Trace Tracking : 차세대공약(Ver.2017) → 차세대공약(Ver.2018)</div> <div>정책/전략적으로 높은 중요도를 갖는 항목으로, 2018년도에도 차세대공약 항목으로 분류</div>						





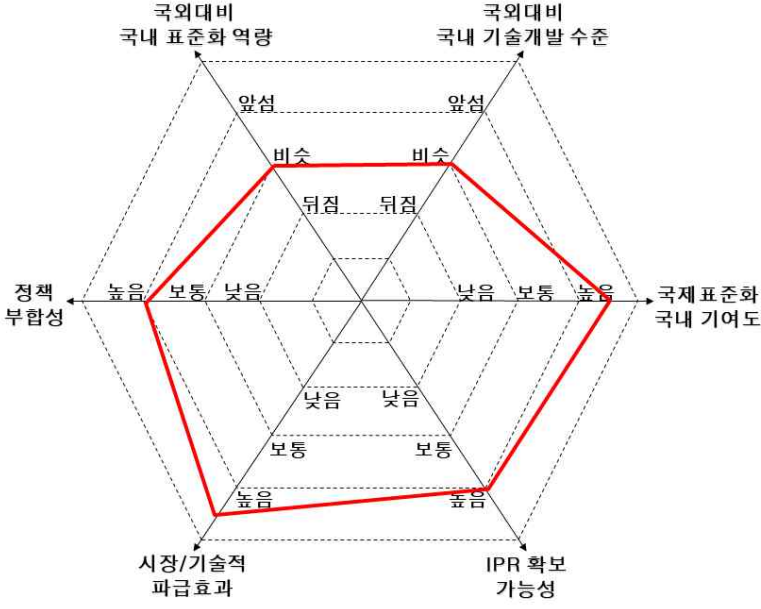
국제 표준화 대응 방안	<p><b>&lt;현황&gt;</b></p> <ul style="list-style-type: none"> <li>ITU-T SG 13에서 2015년 개발이 시작되었으며, 2017년 11월을 목표로 개발이 진행. 우리나라와 중국이 에디터로 표준화를 주도</li> </ul> <p><b>&lt;대응방안&gt;</b></p> <ul style="list-style-type: none"> <li>(국제표준화기구 의장단 수임) 한국이 에디터로서 '빅데이터 교환 프레임워크 및 요구사항 (ITU-T Y.BigDataEX-reqts)' 문서를 2017년 12월 제정을 목표로 표준 개발을 주도하고 관련 표준의 및 국내 데이터 유통 시범사업 등과 연계한 적극적 기고활동이 필요</li> </ul>
국내 표준화 추진 계획	<p><b>&lt;현황&gt;</b></p> <ul style="list-style-type: none"> <li>빅데이터 유통 사업 등 관련 사업을 위한 표준화 요구사항이 논의 중</li> </ul> <p><b>&lt;추진계획&gt;</b></p> <ul style="list-style-type: none"> <li>(국내 표준화 추진) 국내 빅데이터 유통 관련 기술개발 사업들에서의 의견 수렴 결과를 ITU-T를 통해 반영함으로써 우선 국제표준을 개발하고, 이를 TTA SPG22를 통해 국내 표준으로 수용</li> </ul>
표준 특허 전략	<ul style="list-style-type: none"> <li>표준 및 R&amp;D 초중기 전략 : 특허를 통한 표준 아이템 도출 전략</li> <li>(세부전략) 빅데이터 전처리, On-demand 기반 데이터 생성 및 판매, 데이터 전송 효율화 등 데이터 제품 가공 및 전송 관련 IPR 확보</li> </ul>
기술개발 -표준화 -IPR 연계 방안	<ul style="list-style-type: none"> <li>표준화-기술개발 병행추진</li> <li>(세부전략) 사용자 요구사항 기반의 빅데이터 검색, 빅데이터 전송 경량화 및 데이터 연속성 지원기술과 연계하여 IPR 기반의 구현을 통하여 선행 기술을 확보하고, 기 확보된 기술을 표준화 하는 전략이 요구</li> </ul>

## (차세대공략 | 병행) 빅데이터 프로비넌스 요구사항

전략적 중요도 / 국내 역량				표준 화 기구/ 단체	국내	TTA 빅데이터 SPG
					국외	ITU-T SG13
					국내 참여 업체/ 기관	ETRI
기술 개발 단계	국내	□기초연구→■실험→□시작품→□제품화→□사업화		기술 수준	80% (선도국가대비)	
	국외	□기초연구→□실험→□시작품→■제품화→□사업화		기술 격차	2년	
	선도국가/ 기업	인도 / TATA 한국 / ETRI				
표준화 단계	국내	■과제기획→□과제승인→□개발→□검토→□표준채택		표준 수준	90% (선도국가대비)	
	국제	□과제기획→□과제승인→■개발→□검토→□표준채택		표준 격차	1년	
	선도국가 /기업	한국/ETRI, 중국/ZTE				
<p>- Trace Tracking : 차세대공략(Ver.2017) → 차세대공략(Ver.2018)</p> <p>정책/전략적으로 높은 중요도를 갖는 항목으로, 2018년도에도 차세대공략 항목으로 분류</p>						



## (차세대 공략 | 병행) 빅데이터 환경을 위한 메타데이터

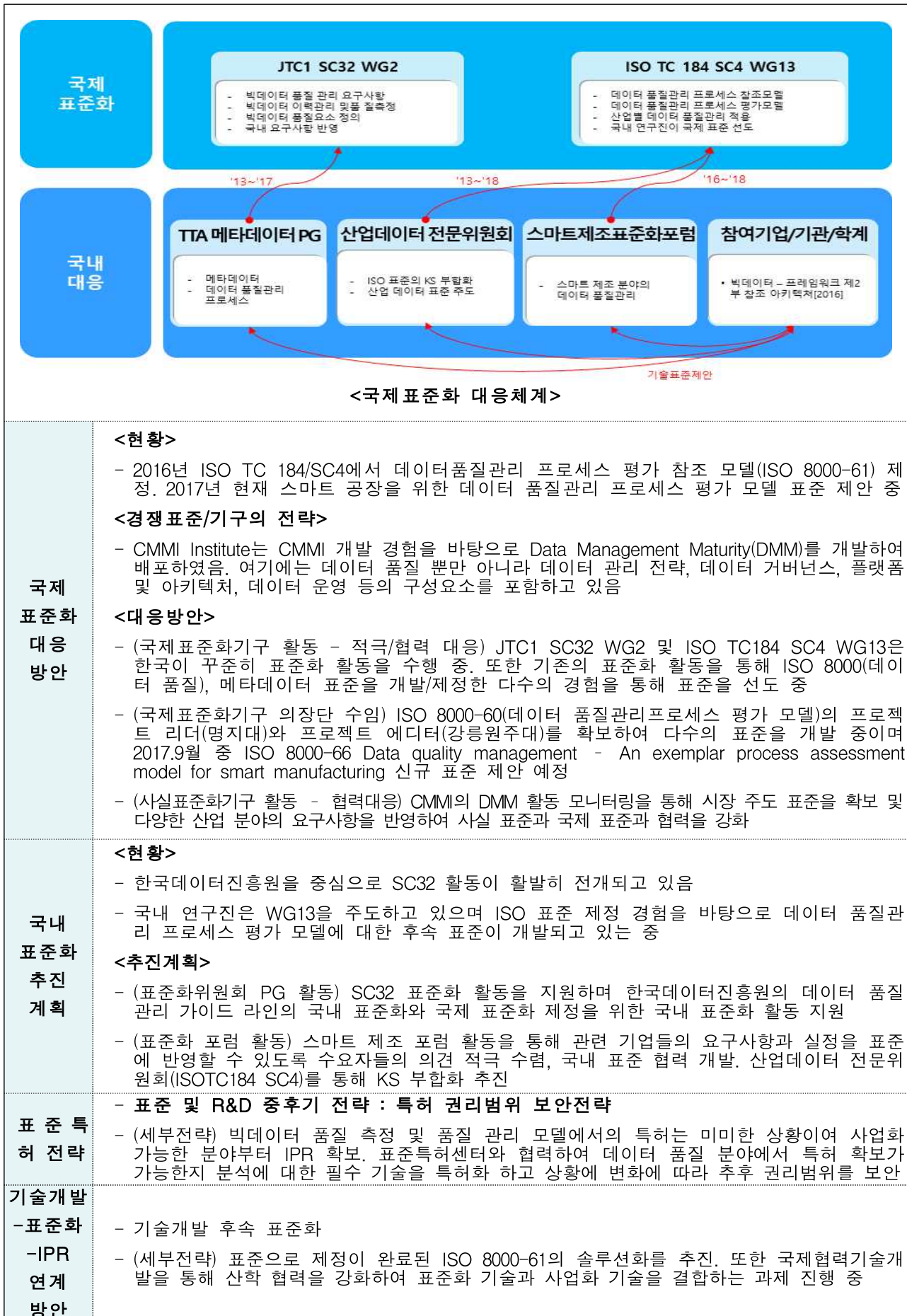
전략적 중요도 / 국내 역량				표준화 기구/ 단체	국내	TTA 메타데이터 PG, 빅데이터 SPG22
	국제	JTC1 WG9, ITU-T SG13				
	국내 참여 업체/ 기관	ETRI				
기술 개발 단계	국내	<input type="checkbox"/> 기초연구→ <input checked="" type="checkbox"/> 실험→ <input type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 수준	90% (선도국가대비)	
	국외	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input checked="" type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 격차	1년	
	선도국가/ 기업	미국 / Apache Foundation, W3C 한국 / 데이터진흥원, ETRI				
표준화 단계	국내	<input checked="" type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 수준	90% (선도국가대비)	
	국제	<input type="checkbox"/> 과제기획→ <input checked="" type="checkbox"/> 과제승인→ <input type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 격차	1년	
	선도국가/ 기업	한국/ETRI, 미국/W3C				
<p>- Trace Tracking : 다각화협력(Ver.2017) → 차세대공략(Ver.2018)</p> <p>빅데이터의 확산에 따른 유통과 통합 관리가 핵심 이슈로 대두되면서 이의 실현기술로 메타데이터 관리에 대한 중요성이 증대되고 있음. 관련 기술 개발을 통한 산업 활성화가 기대되며, 국제표준 제정이 가시화되는 상황이므로 2018년에는 차세대 공략 항목으로 분류함</p>						



## (적극공략 | 후행) 빅데이터 품질측정 및 관리모델

전략적 중요도 / 국내 역량	<p>국제 표준화 국내 기여도</p> <p>정책 부합성</p> <p>시장/기술적 파급효과</p> <p>IPR 확보 가능성</p> <p>국외대비 국내 표준화 역량</p> <p>국외대비 국내 기술개발 수준</p> <p>앞섬</p> <p>비슷</p> <p>뒤짐</p> <p>높음</p> <p>보통</p> <p>낮음</p>			표준화 기구/ 단체	국내	TTA 메타데이터PG, 스마트제조표준 화포럼, 국가기술표준원 ISO TC184 SC4전문위원회
	국외	JTC1 SC32 WG2, ISO TC184 SC4 WG13				
	국내 참여 업체/ 기관	KoDB, 2e컨설팅, 지티원,명지대, 강릉원주대				
기술 개발 단계	국내	□기초연구→□실험→□시작품→■제품화→□사업화			기술 수준	100% (선도국가대비)
	국외	□기초연구→□실험→□시작품→■제품화→□사업화			기술 격차	0년
	선도국가/ 기업	미국 / Google, 아마존, IBM, 오라클 한국 / 다음소프트, SKT				
표준화 단계	국내	□과제기획→□과제승인→□개발→□검토→■표준채택			표준 수준	100% (선도국가대비)
	국제	□과제기획→□과제승인→□개발→□검토→■표준채택			표준 격차	0년
	선도국가 /기업	한국/2e컨설팅, 미국/ECCMA, 독일/Höhn&Partner PartG, 스페인/UCLM				
<p>- Trace Tracking : 방어적수용(Ver.2017) → 적극공략(Ver.2018)</p> <p>국내에서는 민간 데이터 품질 분야에서는 한국데이터진흥원, 공공 데이터 품질 분야에서는 한국정보화진흥원이 주도하고 있으며 전문 기업들이 솔루션 제공과 컨설팅을 수행하고 있음. 해외 솔루션은 데이터 품질 측정과 관리 솔루션 위주로 보급되어 있음. 한국 연구진이 데이터 품질 관련 국제 표준 개발/제정을 선도하고 있으며 국내외에서 국제 표준의 활용도가 높아서 스마트 제조, IoT 등의 전문 분야 적용되고 있으므로 적극공략으로 분류</p>						

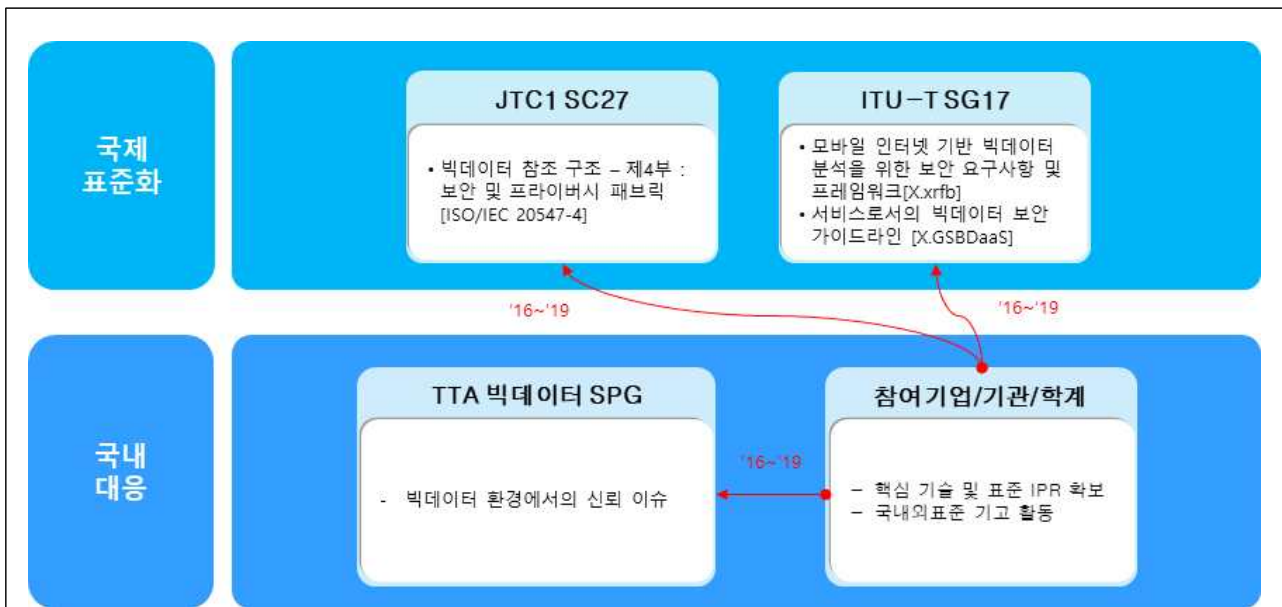






## (차세대공략 | 후행) 빅데이터 프라이버시 및 보안관리

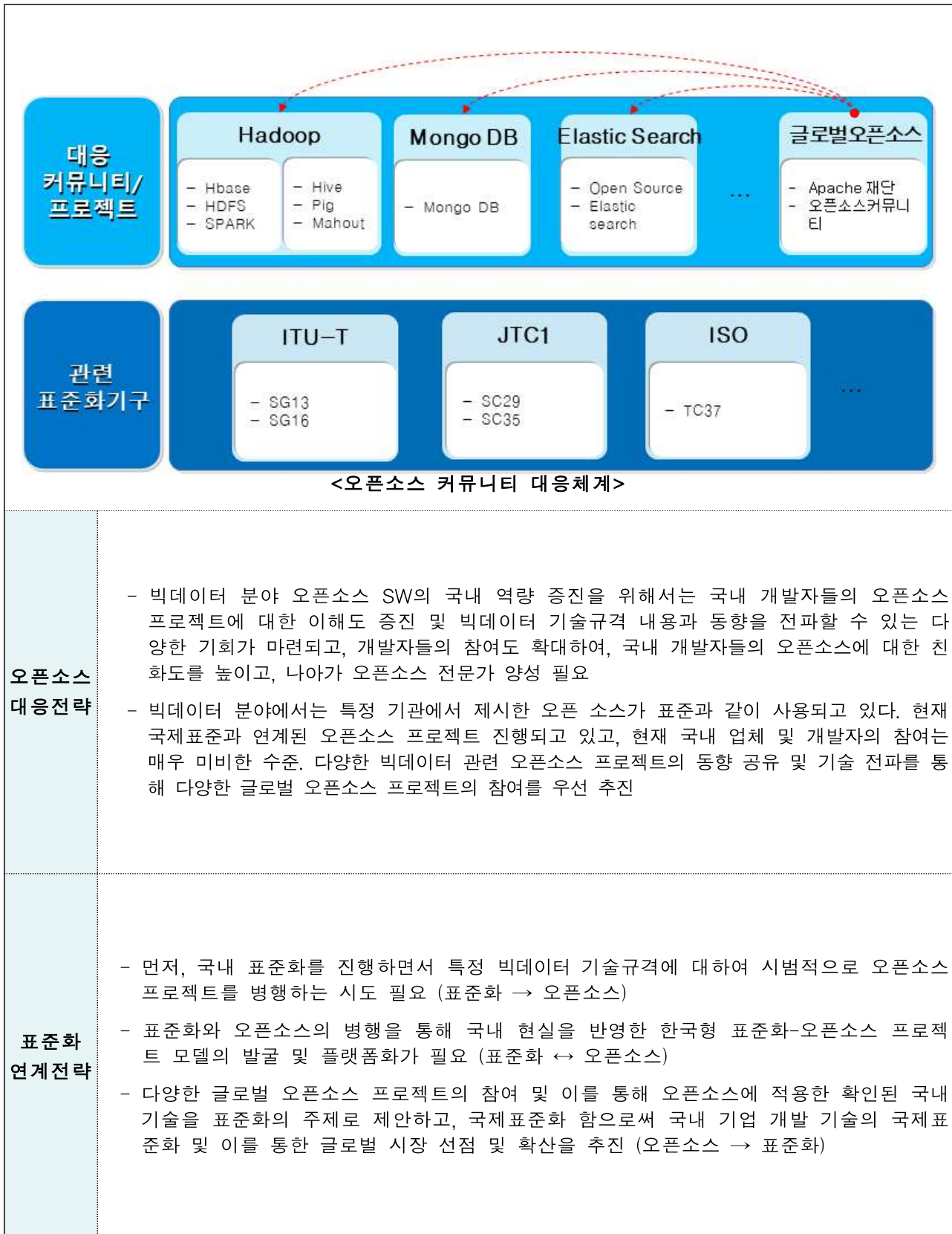
전략적 중요도 / 국내 역량	<p>국외대비 국내 표준화 역량</p> <p>국외대비 국내 기술개발 수준</p> <p>앞섬</p> <p>비슷</p> <p>뒤짐</p> <p>정책 부합성</p> <p>높음</p> <p>보통</p> <p>낮음</p> <p>국제 표준화 국내 기여도</p> <p>시장/기술적 파급효과</p> <p>IPR 확보 가능성</p>			표준화 기구/ 단체	국내	TTA 빅데이터 SPG
	국제	JTC1 SC27, ITU-T SG17				
	국내 참여 업체/ 기관	ETRI, TCA서비스				
기술 개발 단계	국내	<input type="checkbox"/> 기초연구→ <input checked="" type="checkbox"/> 실험→ <input type="checkbox"/> 시작품→ <input type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 수준	75% (선도국가대비)	
	국외	<input type="checkbox"/> 기초연구→ <input type="checkbox"/> 실험→ <input type="checkbox"/> 시작품→ <input checked="" type="checkbox"/> 제품화→ <input type="checkbox"/> 사업화		기술 격차	2.5년	
	선도국가/ 기업	미국 / IBM				
표준화 단계	국내	<input checked="" type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 수준	90% (선도국가대비)	
	국제	<input type="checkbox"/> 과제기획→ <input type="checkbox"/> 과제승인→ <input checked="" type="checkbox"/> 개발→ <input type="checkbox"/> 검토→ <input type="checkbox"/> 표준채택		표준 격차	1.0년	
	선도국가/ 기업	중국 / CECGW 한국 / ETRI				
<p>- Trace Tracking : 차세대공략(Ver.2017) → 차세대공략(Ver.2018)</p> <p>정책/전략적으로 높은 중요도를 갖는 항목으로, 2018년도에도 차세대공략 항목으로 분류</p>						



&lt;국제 표준화 대응체계&gt;

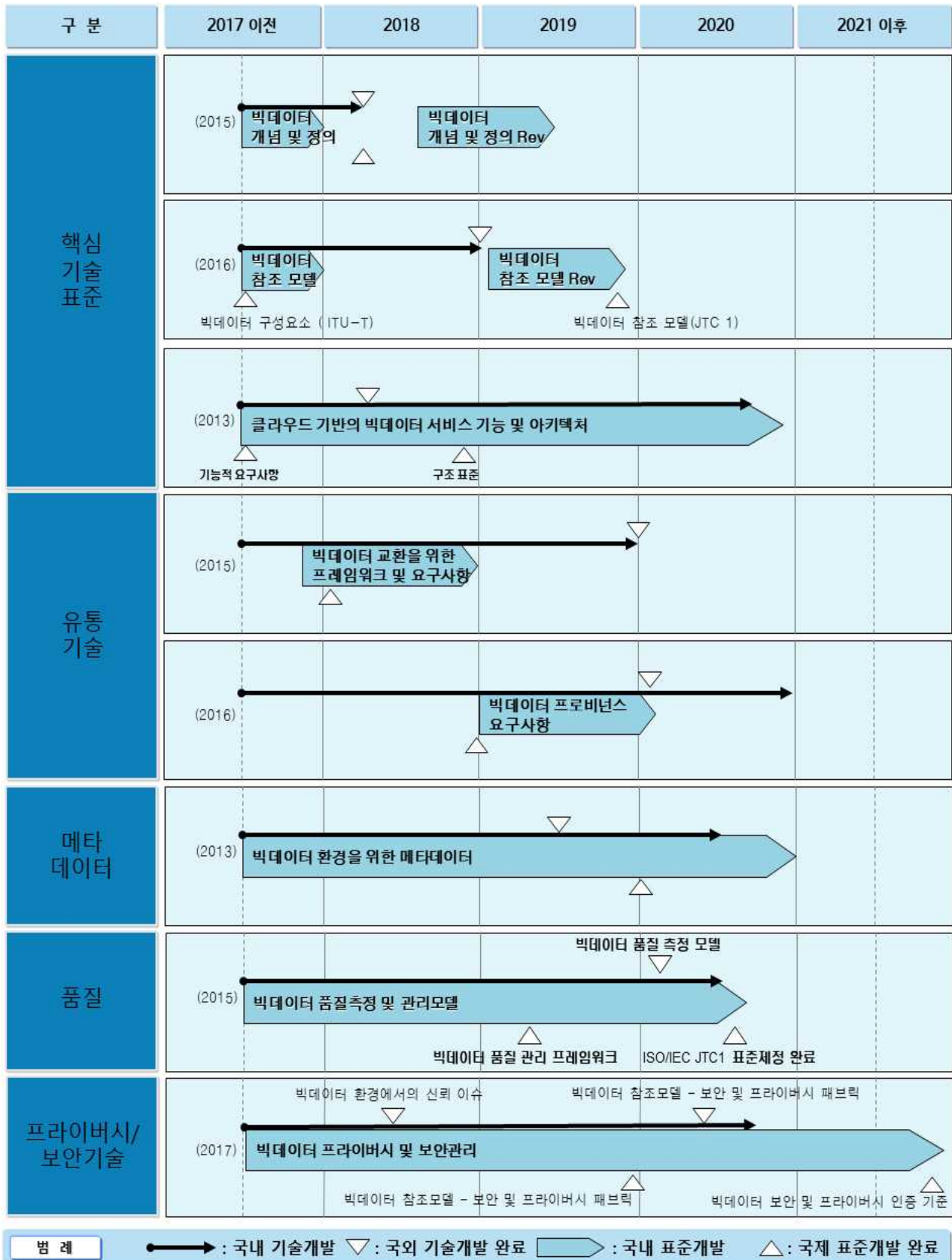
국제 표준화 대응 방안	<p>&lt;현황&gt;</p> <ul style="list-style-type: none"> <li>- ISO/IEC JTC1 WG 9에서 제안된 빅데이터 참조모델의 일부인 4부 보안 및 개인정보보호 패브릭이 ISO/IEC JTC 1 SC 27에서 개발 중에 있음</li> </ul> <p>&lt;경쟁표준/기구의 전략&gt;</p> <ul style="list-style-type: none"> <li>- ITU-T SG17에서는 모바일 인터넷 환경 및 서비스로서의 빅 데이터 등 특정 환경에 기초한 보안 요구사항 및 프레임워크, 가이드라인 등을 개발하고 있음</li> </ul> <p>&lt;대응방안&gt;</p> <ul style="list-style-type: none"> <li>- (국제표준화기구 활동) ISO/IEC JTC 1 SC 27에서 개발 중인 20547-4(빅데이터 참조모델 - 4부 보안 및 개인정보보호 패브릭)는 JTC 1 WG 9에서 한국이 주도하여 수립한 20547의 일부임. 따라서 WG 9과 연계하여 국내 연구 개발된 보안 및 개인정보보호 요구사항을 기고하는 등 적극 대응이 필요함, 또한 ITU-T SG17의 X.srfb(Security requirements and framework for Big Data Analytics in mobile Internet services)는 ETRI가 에디터십을 확보한 상태임, 따라서 이러한 리더십을 이용하여 국내의 기술을 적극적으로 표준화 하는 대응이 필요하며, X.GSDaaS 개발을 주도하고 있는 중국과도 긴밀한 협업 관계를 유지해야 함</li> </ul>
국내 표준화 추진 계획	<p>&lt;현황&gt;</p> <ul style="list-style-type: none"> <li>- TTA SPG 22에서 빅데이터 환경에서의 신뢰 이슈가 표준 2건 개발 중</li> </ul> <p>&lt;추진계획&gt;</p> <ul style="list-style-type: none"> <li>- (국내 표준화 추진) TTA SPG22를 통해 빅데이터 환경에서의 신뢰 이슈를 표준 제정하고, ITU-T 및 JTC1 국제표준 국제표준을 우선 개발하고, 이를 TTA SPG22를 통해 국내 표준으로 수용</li> </ul>
표준 특 히 전략	<ul style="list-style-type: none"> <li>- 표준 초중기 및 R&amp;D 중후기 전략 : 표준 필수특허 설계전략</li> <li>- (세부전략) 현재 국제 표준화 아이템을 구체적으로 도출하고 이를 기반으로 표준화를 진행 하되, 표준화 방향에 따라, 출원 및 표준 기고서 유연하게 작성</li> </ul>
기술개발 -표준화 -IPR 연계 방안	<ul style="list-style-type: none"> <li>- 기술개발 후속 표준화</li> <li>- (세부전략) 관련 R&amp;D 과제 등을 진행함으로써 실효성 등을 검증하고 이를 기반으로 유효적 특허의 개발을 독려하는 것이 필요</li> </ul>

### 3.3. 오픈소스 국내외 추진전략

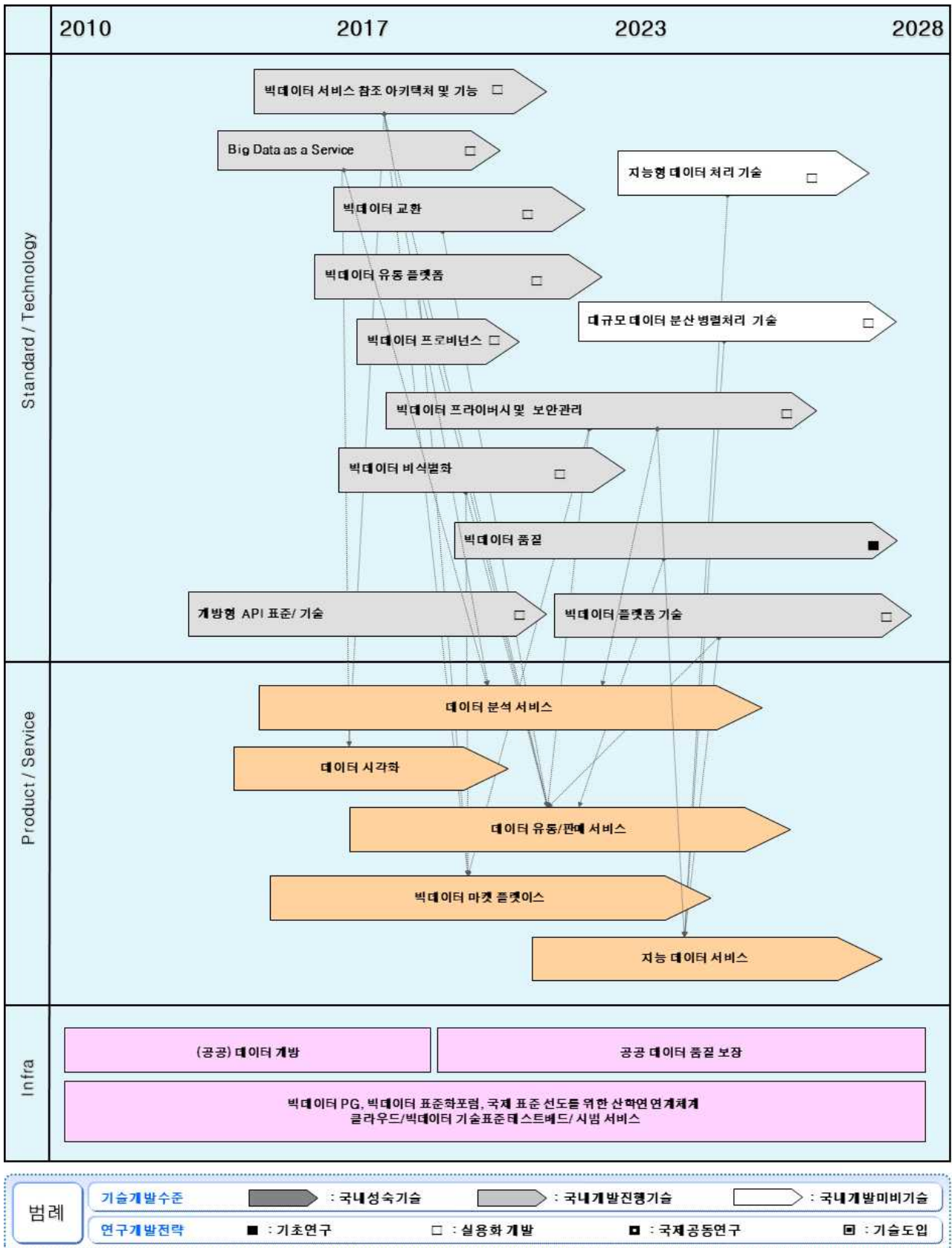


### 3.4. 중기(3개년) 및 장기(10개년) 표준화 계획

#### ○ 중기(2018~2020) 표준화 계획



○ 장기(~2028) 표준화 계획



## [작성위원]

구 분	소 속	성 명	직 위	국 내 외 표 준 화 활 동
총괄	IITP	김형철	CP	▶ 과기정통부 기반SW·컴퓨팅 CP
분과장	숙명여대	윤용익	교수	▶ TTA 빅데이터 (SPG22) 부의장
위원	ETRI	이강찬	책임	▶ ITU-T SG13 Q17 라포처, ISO TC 307 위원 ▶ TTA 클라우드컴퓨팅 (SPG21) 부의장, TTA 빅데이터 (SPG22) 의장
위원	CCCR	김진택	사무국장	▶ TTA 클라우드/빅데이터 특별기술위원회(STC2) 위원
위원	이화여대	양희동	교수	▶ TTA 빅데이터 프로젝트그룹(SP22) 특별위원
위원	TCA 서비스	오경희	대표	▶ ITU-T SG 17 에디터 ▶ TTA 응용보안/평가인증(PG504) 특별위원
위원	강릉원주대학교	이창수	교수	▶ ISO TC 184/SC4/WG13 프로젝트 에디터
위원	ETRI	인민교	책임	▶ ITU-T SG13 Q17/Q18 에디터
위원	KODB	임성준	팀장	▶ TTA 메타데이터 프로젝트그룹(PG606) 의장
자문	금융보안원	임형진	차장	▶ ITU-T WP4 에디터 ▶ TTA 응용보안/평가인증(PG504) 위원
위원	안양대	정의현	교수	▶ TTA IPv6/인터넷주소자원 프로젝트그룹(PG222) 부의장
위원	ETRI	하수욱	책임	▶ ITU-T SG 13 에디터, JTC 1 WG 9 에디터 ▶ TTA 빅데이터 (SPG22) 간사
특허분석	KISTA	박성혁	선임	▶ TTA 표준화전략맵 빅데이터 특허분석
사무국	TTA	박준환	전임	▶ TTA 클라우드컴퓨팅 (SPG21), TTA 빅데이터 (SPG22) 사무국 담당
간사	TTA	김영재	수석	▶ TTA 표준화전략맵 빅데이터 분야 간사



## [참고문헌]

1. Market Watch, "Big Data Equals Big Business opportunity Say Global IT and Business Professionals", 2012.5.14.
2. Informatica, "Balancing Opportunity and Risk in Big Data", 2012.5.15.
3. Y.3600, Initial draft Recommendation Y.BigData-reqts, Requirements and capabilities for cloud computing based big data. ITU-T, 2015.7
4. 정보통신정책연구원, “빅데이터 시장의 현황 및 전망”, 2013.3
5. ISO/IEC JTC 1 SC 32 홈페이지 <http://JTC 1SC 32.org/>
6. 창조경제 및 정부3.0 지원을 위한 빅데이터 산업 발전전략, 미래부 등, 2013.12.
7. 빅데이터 시장·정책·표준화 현황, ETRI 표준연구센터, 2014.2.
8. 한국콘텐츠진흥원, 빅데이터 시장 현황과 콘텐츠산업 분야에 대한 시사점, 2014.1
9. 한국과학기술정보연구원, 빅데이터 산업의 현황과 전망, 2013.4
10. 한국정보화진흥원, 빅데이터 활용 단계별 업무절차 및 기술활용 매뉴얼, 2014.5.
11. 한국정보화진흥원, 알기 쉬운 공공부문 빅데이터 분석·활용 가이드, 2012.12.
12. 방송통신위원회, 빅데이터 개인정보보호 가이드라인, 2014.12
13. 박현아, 빅데이터 시장 현황과 콘텐츠산업 분야에 대한 시사점, 코카포커스, 2013-11호, 한국콘텐츠 진흥원, 2014.1
14. Draft supplement Y.Suppl.BigData-RoadMap, “ITU-T Y.BigData-reqts - Supplement on big data standardization roadmap”, ITU-T, 2015.7
15. 빅데이터 관련 오픈소스 개발동향, 한국통신학회 종합 학술 발표회 논문집, 680-681 (2013)
16. “빅데이터산업의 부상 및 시사점”, 김광섭, 산은조사월보, 2016.10.
17. “IDC 2017년 세계 빅데이터 분석 시장 1500억달러 넘어설 전망”, 조선비즈, 2017. 4. 17, [http://biz.chosun.com/site/data/html\\_dir/2017/04/17/2017041701261.html](http://biz.chosun.com/site/data/html_dir/2017/04/17/2017041701261.html)
18. Recommendation ITU-T Y.3600 (2015), Big data - cloud computing based requirements and capabilities
19. Initial draft Recommendation Y.BDaas-arch, Big data - Functional architecture of Big Data as a Service, [http://www.itu.int/itu-t/workprog/wp\\_item.aspx?isn=10548](http://www.itu.int/itu-t/workprog/wp_item.aspx?isn=10548)
20. Recommendation ITU-T Y.3600 (2015), Big data - cloud computing based requirements and capabilities
21. Initial draft Recommendation Y.BDaas-arch, Big data - Functional architecture of Big Data as a Service, [http://www.itu.int/itu-t/workprog/wp\\_item.aspx?isn=10548](http://www.itu.int/itu-t/workprog/wp_item.aspx?isn=10548)
22. “한국IBM, 기업용 정보보안 신제품군 대거 발표”, <https://www-03.ibm.com/press/kr/ko/pressrelease/39559.wss>, 2012. 11. 19
23. 양경아 외, “빅데이터 시스템의 보안 위협 및 보안 요구사항 분석”, 보안공학연구논문지,



Vol.13, No.6, pp 501-514, 2016. 12. 08

24. 정동원 등, 빅데이터 처리를 위한 메타데이터 표준화에 관한 연구, 한국정보과학회 동계학술발표, 2015.
25. Initial draft Recommendation Y.bdm-sch, Big data - Metadata framework and conceptual model, 2017.
26. W3C Recommendation, Data Catalog Vocabulary (DCAT), 2014.
27. 한국정보화진흥원, 알기 쉬운 Linked Open Data, 2014.

## [약어]

API	Application Programming Interface
ASP	Application Service Provider
BDaaS	Big Data as a Service
CAGR	Compound Annual Growth Rate
CDC	Cloud Data Center
CDG	Cloud Data Governance
CDMI	Cloud Data Management Interface
CG	Community Group
CIF	Cloud Industry Forum
CIMI	Cloud Infrastructure Management Interface
CSA	Cloud Security Alliance
CSB	Cloud Service Brokerage
CSV	Comma-Separated Values
CTP	Cloud Trust Protocol
DaaS	Desktop as a Service
DAP	Device API
DMTF	Distributed Management Task Force
EC2	Elastic Computer Cloud
FGCC	Focus group Cloud Computing
GLD	Government Linked Data
IaaS	Infrastructure as a Service
IETF	Internet Engineering Task Force
ICT	Information and Communications Technologies
ISO	International Organization for Standardization
KIG	Korea Interest Group
KMIP	Key Management Interoperability Protocol
LDP	Linked Data Platform
LOD	Linked Open Data
M2M	Machine-to-Machine
MVNO	Mobile Virtual Network Operator
NaaS	Network as a Service
NAS	Network Attached Storage
NIST	National Institute of Standards and Technology
OASIS	Organization for the Advancement of Structured Information Standards
OCCI	Open Cloud Computing Interface

OCI	Open Container Initiative
OCM	Open Cloud Manifesto
ODCA	Open Data Center Alliance
OGF	Open Grid Forum
OMG	Object Management Group
OSS	Operation support systems
OVF	Open Virtualization Format
OWL	Web Ontology Language
PaaS	Platform as a Service
QoS	Quality of Service
RDF	Resource Description Framework
REST	REpresentational State Transfer
S3	Simple Storage Service
SBC	Server Based Computing
SDO	Standard Developing Organization
SGBD	Study Group on Big Data
SLA	Service Level Agreement
SMI	Storage Management Initiative
SNIA	Storage Networking Industry Association
SOA	Service Oriented Architecture
SPG	Special Project Group
SQL	SQL Query Language
ToR	Term of Reference
VDI	Virtual Desktop Infrastructure
W3C	World Wide Web Consortium
WG	Working Group
WP	Working Party

