

본 강의에서 수업자료로 이용되는 저작물은
저작권법 제25조 수업목적 저작물 이용 보상금제도에 의거,
한국복제전송저작권협회와 약정을 체결하고 적법하게 이용하고 있습니다.
약정범위를 초과하는 사용은 저작권법에 저촉될 수 있으므로
수업자료의 재 복제, 대중 공개·공유 및 수업 목적 외의 사용을 금지합니다.

2021. . .

부천대학교·한국복제전송저작권협회

• 교육 과정 계획

- 01 4차 산업혁명의 개요
- 02 빅데이터 개요
- 03 인공지능 개요
- 04 사물인터넷 개요
- 05 자율주행차 개요
- 06 가상·증강·혼합·확장현실 개요
- 07 드론 개요
- 08 중간고사
- 09 3D프린팅과 헬스케어 개요
- 10 블록체인 개요
- 11 클라우드 컴퓨팅 개요
- 12 신재생에너지와 산업 변화
(또는 산업체직무전문가 특강)
- 13 플랫폼 비즈니스 개요
(또는 산업체직무전문가 특강)
- 14 스마트 생태계 개요
- 15 기말고사



정보처리산업기사 실기
신기술 토픽들

- 학습 목표
 - 빅데이터 개념 익히기
 - 빅데이터 특징과 활용 알아보기
 - 신기술 용어 익히기
- 목차
 - 01 빅데이터 이해
 - 02 빅데이터 특징
 - 03 빅데이터 활용
 - 04 신기술 용어

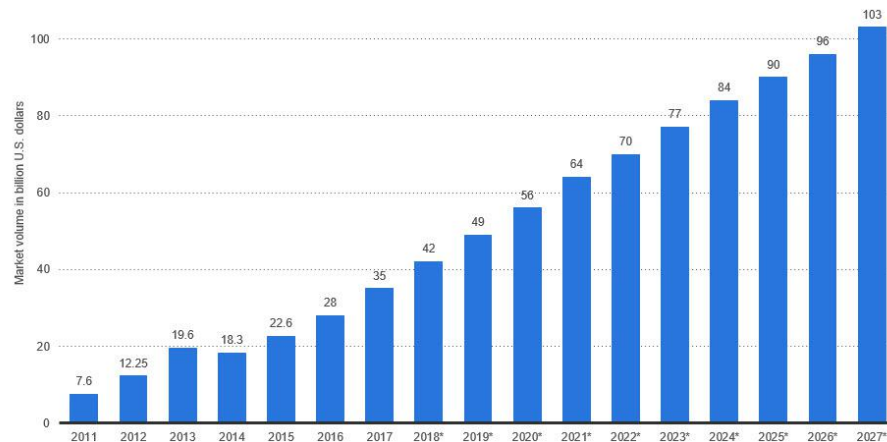
빅데이터 이해

- 빅데이터(Big Data)란?

- 빅데이터 : 기존의 관리 체계 및 분석 방법으로는 해결이 불가능한 막대한 (초대용량) 데이터의 집합
- 규모 방대, 생성 기간 짧음, 정형·비정형 데이터를 총괄하는 대규모 형태

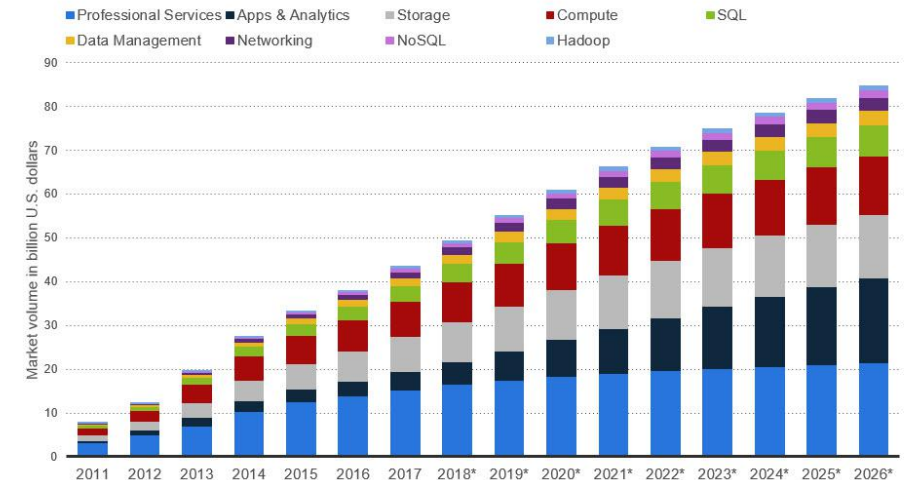
Forecast Revenue Big Data Market Worldwide 2011-2027

Big Data Market Size Revenue Forecast Worldwide From 2011 To 2027
(in billion U.S. dollars)



Big Data Market Worldwide Segment Revenue Forecast 2011-2026

Big Data Market Forecast Worldwide from 2011 to 2026, by segment
(in billion U.S. dollars)



[이미지출처] Wikibon and reported by Statist,

<https://blogs-images.forbes.com/louiscolumbus/files/2018/05/Big-Data-Market-Size-Revenue-Forecast-Worldwide-From-2011-To-2027-Final.jpg>

<https://blogs-images.forbes.com/louiscolumbus/files/2018/05/Big-Data-Market-Forecast-Worldwide-By-Segment.jpg>

빅데이터 이해

- 정보화 세계의 패러다임의 흐름
 - 스마트 혁명으로 데이터 폭증 -> 빅데이터 출현 -> 데이터 패러다임 변화
 - 정보 사회를 움직이는 데 있어서 데이터는 필수 요소
 - 빅데이터 고급 분석 관련 기술은 현재 기술 발생단계

구분	PC환경	인터넷 환경	모바일 환경	스마트폰 환경
시대별	디지털 시대	온라인 시대	모바일 시대	사물지능 시대
주요 환경	전산, PC통신	웹네트워크(WWW), 웹2.0	스마트기기, 모바일 인터넷	인공지능(AOI), 사물인터넷(IoT)
핵심 서비스	운영체제(OS), 응용프로그램(APP)	포털(검색엔진), 전자상거래(EC)	SNS 플랫폼, 스마트폰	무인자동차, 로봇 시스템
대표기업	마이크로소프트, IBM 등	네이버, 다음, 야후 등	애플, 페이스북, 트위터, 카카오 등	구글, 아마존, 삼성 등

* 웹2.0(Web 2.0) : 공유, 개방, 참여를 바탕으로 쌍방향 소통 웹 기술

- 빅데이터 정의들

- 위키피디아 정의

- : 기존 데이터베이스 시스템의 데이터 수집-> 저장 -> 관리 -> 분석 기능을 넘어서는 정형·비정형 데이터 세트로부터 가치를 추출하고 결과를 분석하는 기술

- 국가전략위원회 정의

- : 대용량 데이터에서 가치 있는 정보를 추출하고 생성된 지식을 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술

- 삼성경제연구소 정의

- : 빅데이터란 기존의 관리 및 분석 체계로는 감당할 수 없을 정도의 거대한 데이터 집합으로 대규모 데이터와 관계된 기술 및 도구(수집·저장·검색·공유·분석·시각화 등)을 모두 포함하는 개념

→ (양적 + 질적) 데이터로서 의미를 지님

- 빅데이터 구성

- 빅데이터는 정형화 데이터, 반정형화 데이터, 비정형화 데이터 등이 복합적으로 이루어짐
- 정형화 데이터 : 일정한 규칙에 따라 체계적으로 정리된 데이터(통계자료)
- 반정형화 데이터 : 문서 프로그램 등의 응용 소프트웨어로 작성된 데이터
- 비정형화 데이터 : 동영상, 음악, 카카오톡 등 공유 및 정보교환이 주목

구분	기존 데이터	빅데이터
데이터	- 정형화된 수치자료(통계자료)	- 비정형 데이터(SNS, 모바일 자료)
하드웨어	- 파일 저장장치 - 데이터베이스	- 클라우드 컴퓨팅(컴퓨터 자원의 효율적인 제공 활용)
소프트웨어	- 관계형 데이터베이스(SQL) - 통계SW 툴 - 데이터마이닝	- 개방형 소스 - 버즈분석 및 감성분석 - 텍스트마이닝

- 데이터 분석

- 버즈분석(buzz analysis) : 온라인에 떠도는 게시물 분석
- 사실(Fact) : 텍스트 마이닝
- 태도(Attitude) : 오피니언 마이닝
- 감성분석(sentiment analysis) : 오피니언 마이닝(Opinion Mining)이라고도 함, 텍스트에서 감성, 태도 등 주관적 정보를 분석
- 데이터마이닝(data mining) : 데이터 중에서 의미 있는 유용한 정보를 찾아내는 분석
- 텍스트마이닝(text mining) : 텍스트 데이터에서 주제를 찾아 연관 관계 등을 분석
- 웹마이닝(web mining) : 웹사이트 데이터에서 유용한 정보를 찾아내는 분석
- 소셜마이닝(social mining) : 소셜 미디어 글과 활동 등을 분석
- 현실마이닝(reality mining) : 스마트폰 등 현실에서 발생하는 라이프 스타일 정보 등을 분석

- 빅데이터 수집, 저장, 분석

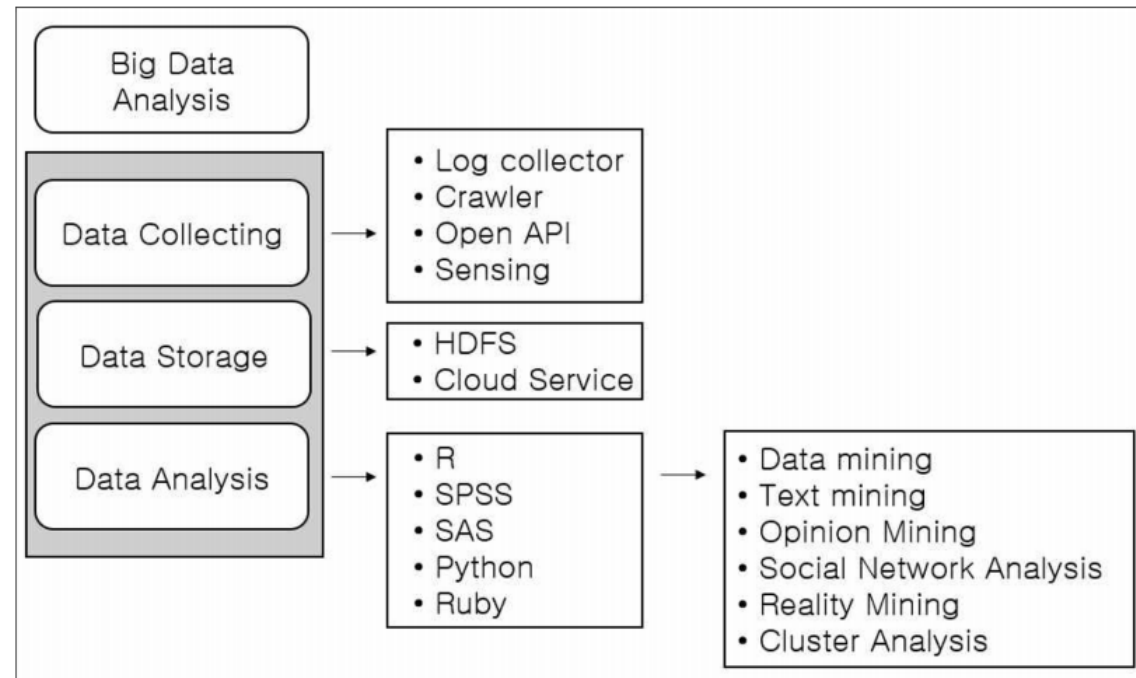
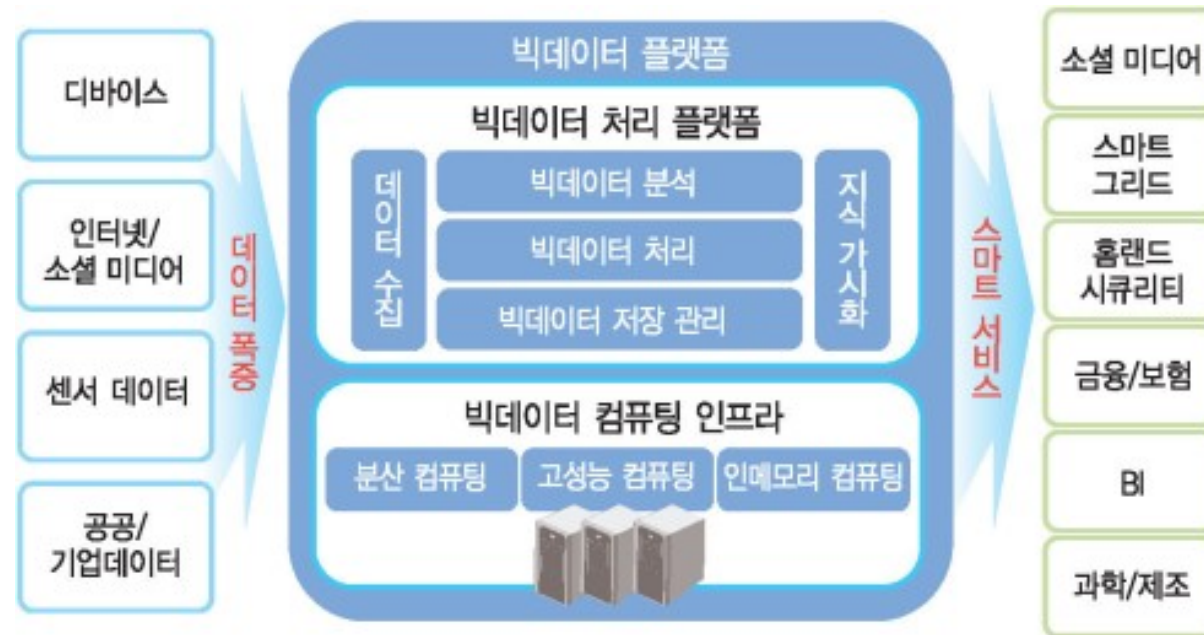


Fig. 1. Summary of big data analysis.

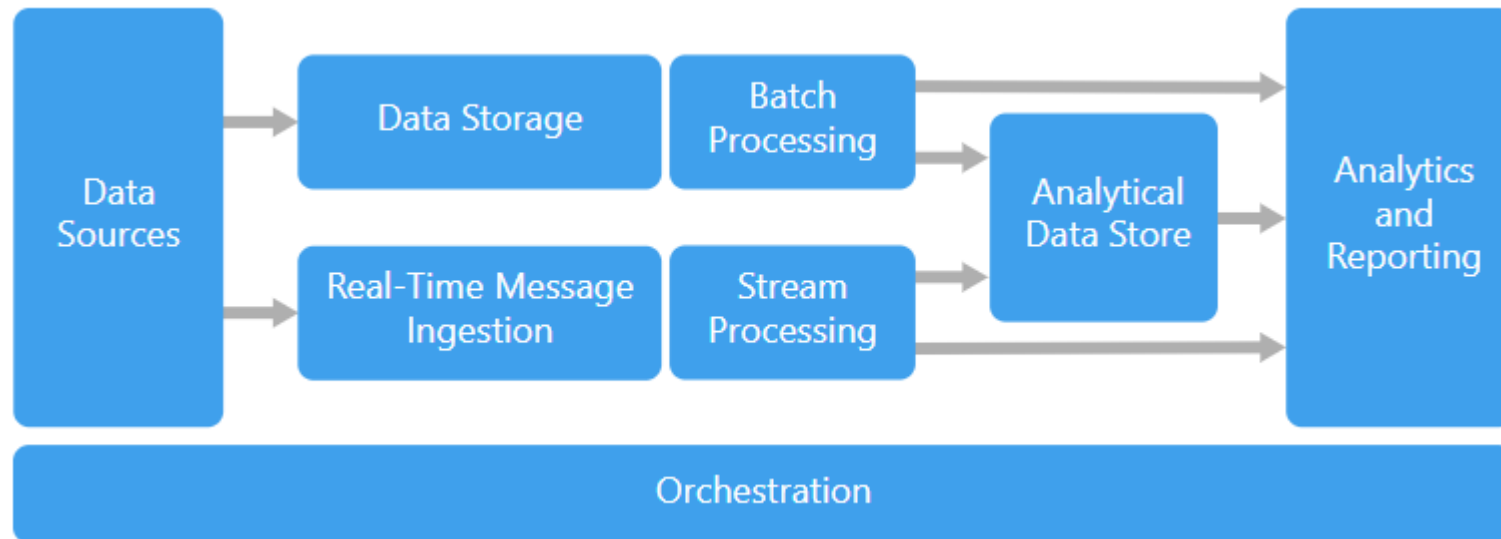
빅데이터 이해

- 빅데이터 플랫폼

- 대용량의 데이터를 분석·처리·관리할 수 있는 컴퓨팅 인프라를 갖추어 폭증하는 데이터를 고객이 원하는 스마트한 데이터로 서비스해주는 구조



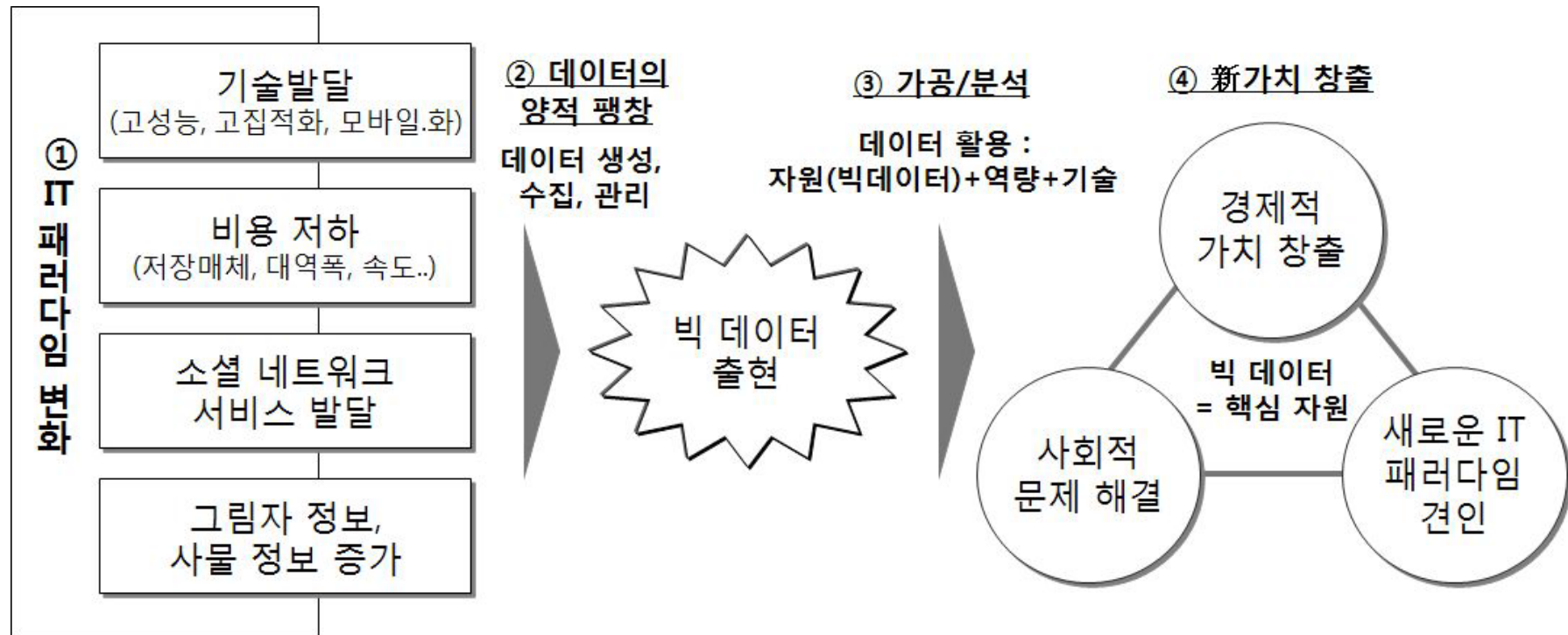
- 빅데이터 아키텍처
 - 데이터 수집, 저장, 처리, 분석, 시각화 단계를 논리적으로 구성



* 오케스트레이션 : 프로세스 최적화 및 자동화를 의미

빅데이터 특징

- 빅데이터 활성화 요인 3가지
 - 스마트폰 보급의 활성화 : 비정형 데이터 폭발적 증가
 - 클라우드 서비스 : 사용자 데이터 축적 -> 분석 -> 재활용 요구 증대
 - SNS 등의 소셜 미디어의 일상화 : 정보 유통 구조 새롭게 재편(쌍방교류)



빅데이터 특징

- 빅데이터 특성과 내용
 - 3V : 용량성(Volume), 다양성(Variety), 신속성(Velocity)
 - 5V : 3V + 정확성(Veracity, 신뢰정도), 가치성(Value)
 - 6V : 5V + 시각화(Visualization)
 - 7V : 5V + 신뢰성(Validity, 타당정도), 휘발성(Volatility)

구분	주요 내용
용량성(volume)	기하급수적으로 증가하는 디지털 정보량 -> 제타바이트(ZB, 1조1000억 기가바이트 (GB)) 시대
다양성(variety)	비정형화된 데이터의 증가 -> SNS, GPS, IoT, AI 데이터 등
신속성(velocity)	실시간성 정보 요구 및 증대 -> 사물 정보(센서, 모니터링), 스트리밍 정보, 모바일 정보의 실시간 처리
정확성(Veracity)	정밀한 정보 처리를 위한 빅데이터 분석 -> 무인화와 로봇 시스템의 질 높은 데이터 필요
가치성(Value)	창조와 가치 창출 -> 기업과 사회의 문제 해결을 위한 유용한 가치를 제공

빅데이터 활용

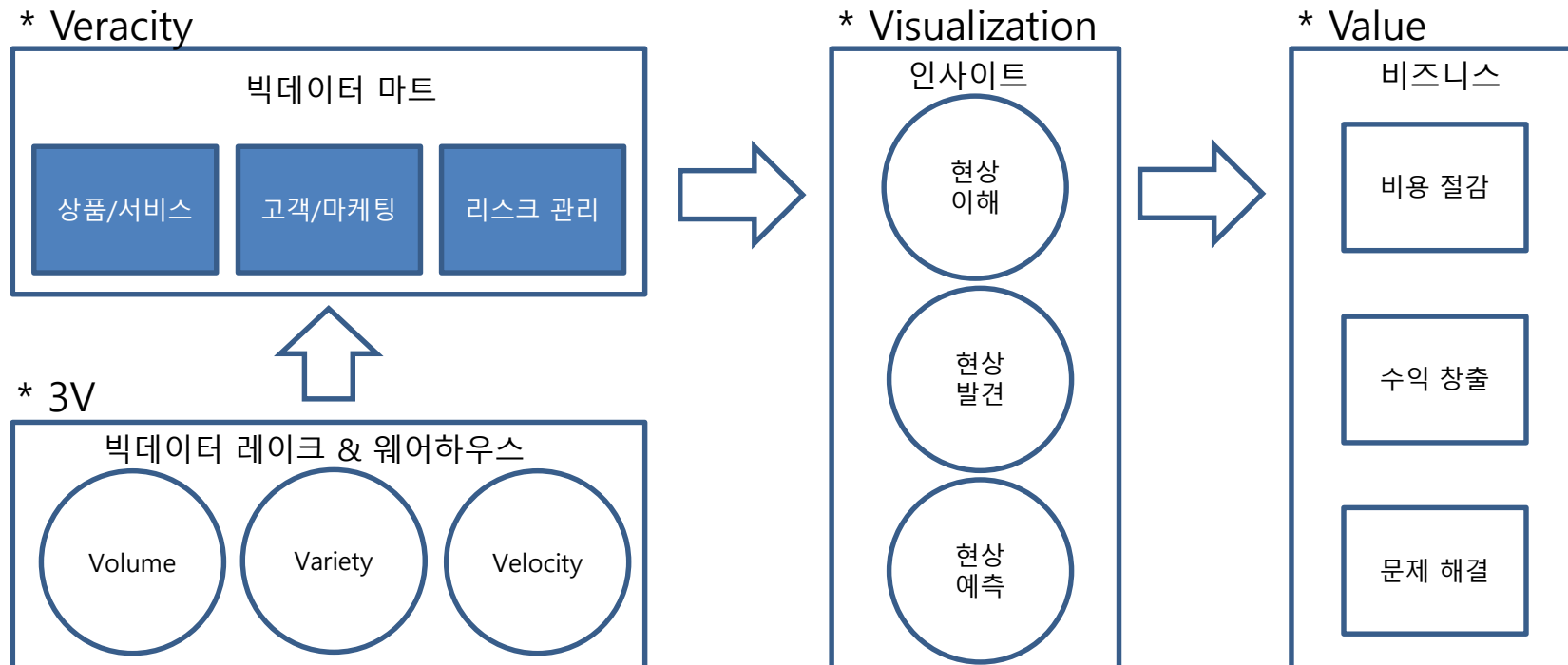
- 빅데이터 활용 사례 – 공공 빅데이터 우수사례집(행정안전부)



1 재난안전	8 문화관광
골드타임 확보를 위한 구급차 배치·운영 최적화 모델 - 전라북도	전주 한옥마을 관광분석을 통한 경제활성화 - 전주시
2 시민안전	9 문화관광
OCTV 설치지역 분석 및 모델 고도화 - 경기도	빅데이터 활용 문화·관광 축제 분석 - 한국관광공사
3 국가지안	10 기상환경농업
인공지능을 활용한 과학적 민생치안강화(임정일지 기반 유사사건 여파분석) - 경찰청·국가정보자원관리원	기상데이터와 농산물 생산성 예측 - 기상청, 농촌진흥청
4 보건의료	11 고용노동
국민건강알람서비스 - 건강보험공단	근로감독 사업장 선정 과학화 - 고용노동부
5 국토교통	12 국토주택
빅데이터 분석을 통한 시내버스 효율적 운영 - 광주광역시	공동주택관리비 빅데이터 분석 - 경기도, 국토교통부
6 도시환경	13 도시안전
전기차 충전 인프라 설치 입지선정 - 대구광역시	인공지능(딥러닝) 기반의 도로포장파손 실시간 탐지시스템 개발 및 적용 - 한국도로공사 ICT센터
7 보건복지	14 공공행정
잠재적 사회취약계층 일자리 창출 및 자립지원 - 남양주시, 국민연금공단	민원분석을 통한 갈등문제 해결 - 포항시

- 빅데이터 목적 및 활용 방안[3]

- 빅데이터 목적 : 비용 절감, 수익 창출, 문제 해결
- 빅데이터 인사이트(통찰력) 3가지 유형 : 이해, 발견, 예측
 - 현상 이해 : 과거 발생한 일에 대한 이해와 원리 파악
 - 현상 발견 : 새로운 패턴 발견과 해석을 통해 무슨 일이 새롭게 일어났는지 파악
 - 현상 예측 : 이해와 발견을 기반으로 예측 모델을 만들어 미래에 발생할 현상 예측



- 국내 업종별 정책 현황
 - 경제 부문
 - 손실을 줄이고 생산성을 높이는데 활용
 - 소비자의 패턴을 분석하여 시장의 동향을 예측
 - 공공 부문
 - 기상 데이터, 인구 데이터 등을 토대로 미래 전략 수립
 - 사회 부문
 - 사회적 약자 도움과 새로운 활동 방안 제시(심야 버스 노선 운영 등)

- 해외 정책 현황

- 미국

- 빅데이터 활용에 관해 행정명령 형태이 가이드라인으로 범정부 전략 수립
 - 빅데이터 교육 및 훈련 환경 개선 강조-> 데이터 과학자 양성 등
 - 프라이버시 강화 기술과 자료 보안 기술 개선

- 일본

- 액티브 데이터 전략 수립 이후 빅데이터를 핵심 요소로 선정
 - 공공 데이터 플랫폼 운영(17개 분야 2만5,000개의 데이터 세트 등재)

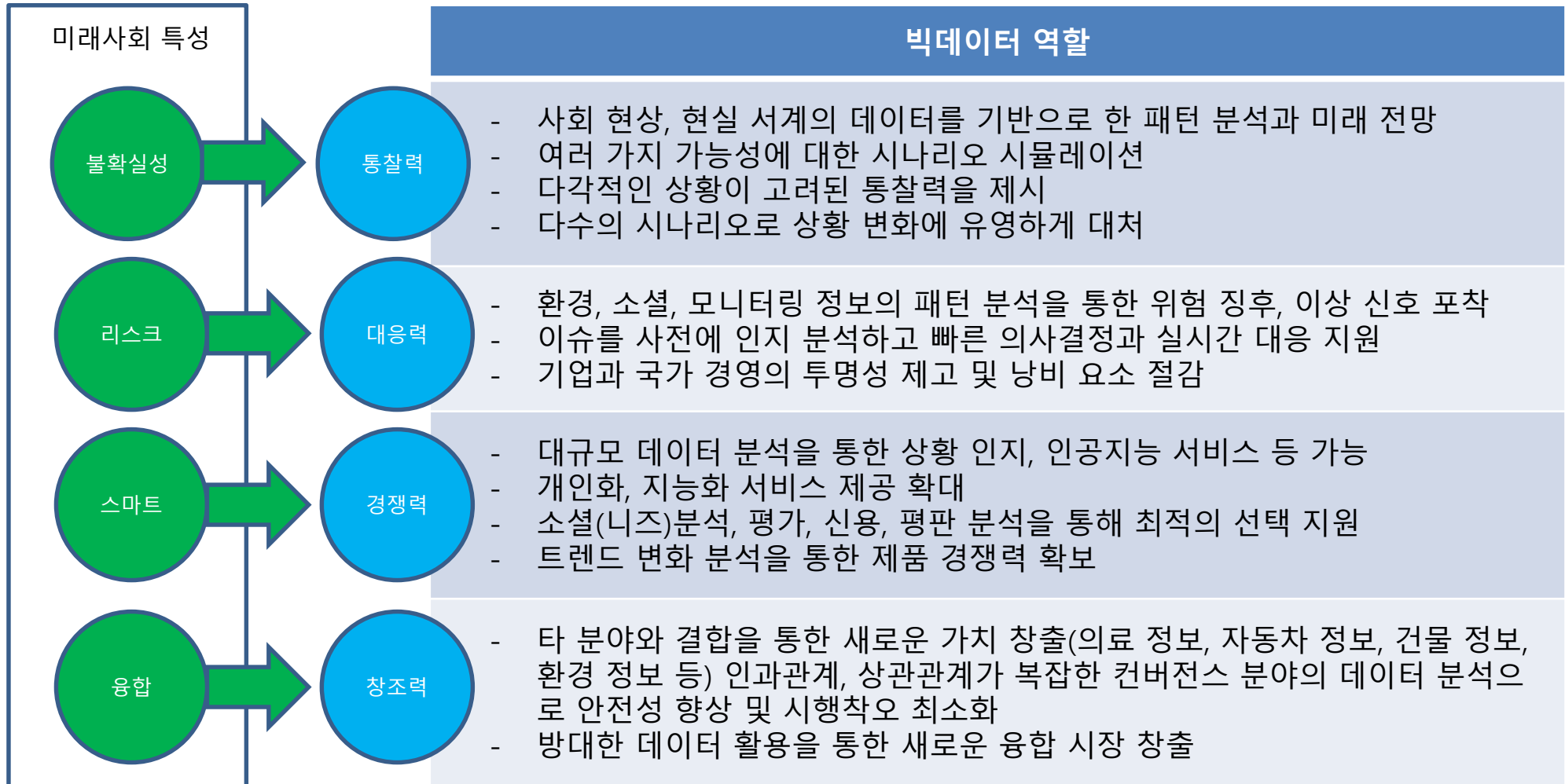
- 중국

- 빅데이터 산업을 국가 신산업으로 지정 -> 적극적인 지원 정책으로 빅데이터 산업 생태계가 조성됨

- 영국

- 데이터산업 활성화 국가 중 하나로 적극적 데이터 관련 정책 수립
 - 공공데이터 포털사이트 운영(5만 1,173개 데이터 세트 등재)
 - 데이터 전문인력 양성을 위해 6-16세 대상 컴퓨팅을 교육과정에 반영

• 미래 사회의 특성과 빅데이터의 역할



- 성공적인 빅데이터 활용 전략

- * 빅데이터 수용 4단계 이론 – 교육, 탐색, 시험, 실행

- 교육

- 지식기반 형성하고 관련된 지식 수집에 중점을 두는 단계
 - 빅데이터 개념 및 정의에 대한 이해 제시, 관련 시장 탐색 및 관측 활동 진행

- 탐색

- 기업 요구와 도전 과제에 기반을 두고 전략과 로드맵을 개발하는 단계
 - 데이터 검색 후 시각화하여 결과에 대한 합리적 이해를 도출하고 의사결정에 반영

- 시험

- 빅데이터를 시범적으로 수용하고 가치와 요구사항을 검증하기 위한 단계
 - 빅데이터 파일럿 프로젝트를 진행하고 가치와 유효성을 검증

- 실행

- 빅데이터를 기업 업무에 적용하는 단계
 - 실제 적용함에 있어 통찰력을 강화하고 피드백을 중점적으로 고려함

- 성공적인 빅데이터 활용 전략
 - 빅데이터 성공적인 활용을 위한 3대 요소 - **자원, 기술, 인력**

<div>자원</div> <div>빅데이터</div>	<div>기술</div> <div>빅데이터 플랫폼</div>	<div>인력</div> <div>빅데이터 사이언티스트</div>
<ul style="list-style-type: none"> - 빅데이터 자원 확보 - 데이터 품질 관리 	<ul style="list-style-type: none"> - 데이터 저장 관리기술(NoSQL, ETL 등) - 대용량 데이터 처리(Hadoop, MapReduce 등) - 빅데이터 분석(자연어 처리, 의미 분석, 데이터 마이닝 등) 시각화 	<ul style="list-style-type: none"> - 수학, 공학 능력 - 경제학, 통계학, 심리학 등 다문화적 이해 - 비판적 시각과 커뮤니케이션 능력 - 스토리텔링 등 시각화 능력

빅데이터 활용

• 빅데이터 분야 기술 범위

〈표 2〉 혁신성장동력 빅데이터 분야 기술 범위

기술분류	핵심기술	기술정의
플랫폼	데이터 자가 증식 및 수집·정제 기술	데이터의 양적인 확대를 위하여 데이터 증식 알고리즘을 활용한 자가 증식을 하거나, 유효하지 않은 데이터를 필터링하거나 샘플링, 정제, 수집하는 기술
	다양한 응용 패턴 통합 지원 기술	데이터가 실제 사용되는 시점에 데이터 사용 목적에 따른 데이터 모델에 맞추어 실시간으로 데이터를 구성하여 제공하고, 다양한 응용 패턴 (배치, 대화형, 스트림 등)을 통합하여 동시 수행을 지원하는 멀티 타입 빅데이터 처리 프레임워크
	멀티모델 데이터 통합, 고신뢰 데이터 관리 및 다각도 분석 기술	분석 목적에 맞게 다양한 모델의 데이터를 통합하고, 데이터의 신뢰성을 확보하면서 통계적으로 중요도를 갖는 결과의 자동 탐색 및 실시간 다각도 분석 기술
	초연결 데이터 관리 및 협업 기술	초연결 인공지능 구현을 위하여 물리적인 데이터 위치나 종류와 무관하게 데이터를 제공할 수 있는 초연결 데이터를 관리하고, 최적의 분석 결과 도출을 위한 다수의 다양한 지능 객체간 집단 협업 지능 플랫폼 기술
	빅데이터 처리 및 저장·관리 기술	질의 처리 성능 향상을 위해 성능 가속 HW, 통신 가속, 인메모리 컴퓨팅 기술을 활용하는 고속 빅데이터 처리 및 저장·관리 기술
분석	지능형 예측 분석 기술	데이터에 숨겨진 패턴을 찾아 과거와 현재의 상황 이해를 바탕으로 미래상황을 예측함으로써, 선제적인 의사결정을 지원
	이종 소스 심층 융합 분석 기술	비정형 텍스트, 관계형 DB 저장 데이터와 더불어 이미지/비디오 및 IoT 스트림 데이터 등 복합형 데이터를 대상으로 통합 분석
	엣지 분석 및 협업 분석 기술	초연결 시대에 발생하는 패스트 데이터에 대한 엣지 분석과 영역별로 산재하는 다수의 엣지 분석 플랫폼들이 연계하여 하나의 글로벌 문제를 분석하고 해결하는 분산/협업형 데이터 분석 기술
	모사현실 모델링 프레임워크	복잡한 실세계를 모사현실로 구현하는 대규모 개방형 모델링 프레임워크 및 최적화 기술
활용	빅데이터 유통 플랫폼 기술	공공·민간의 자유롭고 편리한 데이터 등록, 검색, 활용을 지원하는 플랫폼과 데이터 익명화와 같은 개인정보 보안성을 제공하는 빅데이터 유통 인프라 구축
	워크플로우 기반 적용 시나리오 구현 기술	빅데이터 플랫폼의 적용 범위 확대를 위해 응용 분야별 특화된 적용 시나리오를 워크플로우 기반으로 제공함으로써 빅데이터 플랫폼 및 분석 기술의 활용성 제고
	데이터 품질 정량화 및 최적화 기술	데이터 가치 향상(Value-up)을 위한 데이터의 체계적 축적 및 지속적 관리 체계를 구축하는 데이터 라이프 사이클 관리 기술과 데이터의 품질 진단 및 개선 기술
	빅데이터 응용·서비스 기술	누적된 데이터 또는 실시간 데이터를 발생시키는 다양한 산업분야(의료·건강, 소비·거래, 에너지, 재난안전 분야 등)의 Domain Knowledge와 융합하여 빅데이터 플랫폼 및 분석 기술을 적용, 활용하는 응용·서비스 기술

[함께 생각해 봅시다]
대중교통 문제점과 데이터 기반 해결 방안(아이디어)

- 현재 문제점 : 교통체증 및 대기시간, 노선 부족, 환승, 교통비 등
- 해결책 : 도로 개선 사업, 철도 개선 사업 등

빅데이터 기반 해결책을 함께 생각해 봅시다!

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - VPN(Virtual Private Network, 가설사설망)
 - 사용자가 마치 자신의 전용회선처럼 사용할 수 있도록 해주는 서비스
 - 인터넷망을 전용선의 사설망을 구축하는 것처럼 이용하여 지역적인 제한없이 업무 수행이 가능
 - 터널링 기술 : SSL VPN, IPSEC VPN, PPTP VPN
 - 무선랜(WLAN, Wireless LAN)
 - AP가 설치된 곳에서 RF(무선 주파수)기술을 이용하여 근거리에서 WLAN 카드가 장착된 개인 휴대정보 단말기를 이용해 인터넷에 연결하는 네트워크
 - 802.11 표준기술
 - 직교 주파수 분할 다중(OFDM, Orthogonal Frequency Division Multiplexing)
 - 고속의 데이터를 각 반송파가 직교 관계에 있는 다수의 부반송파에 나누어 실어 다중 전송하는 디지털 변조 방식

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - FTP(File Transfer Protocol, 파일 전송 프로토콜)
 - 파일을 주고 받을 수 있도록 해주는 원격 파일 전송 프로토콜
 - 수동모드(Passive mode)와 능동모드(Active mode, 신호명령 21번/데이터 20번포트) 지원
 - Protocol(프로토콜)
 - 서로 다른 컴퓨터 사이 정보 교환할 수 있게 해주는 표준화된 통신 규약
 - 다중 경로 페이딩(Multipath Fading)
 - 전파가 장애물을 만나 반사 등으로 분산되고 분산된 2개 이상의 신호들이 서로 다른 경로를 통해 수신기에 도달하는 현상(수신 장애 발생)
 - 방지책 : 직교 주파수 분할 다중(OFDM), 다이버시티, 이퀄라이저
 - 원격 검침(AMR; Automatic Meter Reading, Remote Meter Reading Service)
 - 수도 계량기 등 검침원이 고객을 방문하지 않고 원격에서 단말기를 이용해 검침 데이터를 읽을 수 있는 시스템
 - 전화선이나 전력선, 무선주파수 방식으로 정보 수집

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - NGN(Next Generation Network, 차세대 통신망)
 - ITU-T에서 개발하고 있는 유선망 기반의 차세대 통신망
 - 유선 접속망 뿐만아니라 이동 사용자 지원까지를 목표로 하며, 이동통신에서 제공하는 완전한 이동성(Full Mobility) 제공을 목표로 개발
 - GLONASS(글로나스)
 - 미국의 GPS와 유사한 러시아 전파 위성 항법 시스템
 - GPS는 CDMA 방식, GLONASS는 FDMA 방식
 - (참고) GNSS : 미국 GPS, 러시아 GLONASS, 유럽 Galileo, 중국 Beidou 등 모든 위성 항법 시스템 통칭
 - MICS(Medical Implant Communication Service)
 - WBAN의 의료 분야에서 임플란트 장치 간의 통신 서비스 명칭
 - 주파수 402~405MHz를 사용하고, 대역폭은 300KHz

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - SDN(Software Defined Networking, 소프트웨어 정의 네트워킹)
 - 네트워킹을 가상화 하여 제어하고 관리하는 기술
 - 네트워크 비용 및 복잡성을 해결할 수 있는 기술로 네트워크 제어부와 데이터 전달부 인터페이스는 오픈플로(OpenFlow)라는 프로토콜을 활용
 - 근거리 무선 통신(NFC, Near Field Communication)
 - 고주파(HF)를 이용한 근거리 무선통신 기술
 - 아주 가까운 거리에서 양방향 통신을 지원하는 RFID 기술의 일종
 - 13.56MHz 주파수 이용 10cm안에서 최고 424kbps의 속도로 데이터 전송을 지원
 - QoS(Quality of Service, 서비스 품질)
 - 서비스의 품질을 의미하는 것으로 네트워크에서 일정 기준 이하의 지연시간이나 데이터 손실률 등을 보증하기 위한 서비스 규격

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - MODEM(모뎀)
 - 디지털 신호를 아날로그 신호로 변환하는 변조기(modulator), 아날로그 신호를 디지털 신호로 변환하는 복조기(demodulator)를 복합한 장치
 - TEIN(Trans-Eurasia Information Network, 트랜스 유라시아 네트워크)
 - 아시아와 유럽을 연결하는 초고속 정보 통신망
 - 아시아와 유럽 대륙 국가들에게 경제 발전과 교류를 이루는 데 목적
 - Carrier Aggregation(반송파 묶음)
 - 복수의 주파수 대역을 동시에 사용할 수 있는 대역폭 확장 기술(묶어서 전송하는 기술)
 - LTE-Advanced에서 20MHz 대역폭을 갖는 요소 반송파(Component Carrier)를 최대 5개까지 묶어 100MHz 확장된 대역폭을 사용

- 정보처리산업기사(기사) 실기 시험 대비 신기술 용어 익히기[2,4]
 - MAN(Metropolitan Area Network, 도시권 통신망)
 - LAN과 WAN의 중간 형태로 LAN의 기능을 충분히 수용하면서 도시 전역 또는 도시와 도시 등 넓은 지역을 연결하는 통신망
 - VOD(Video On Demand, 주문형 비디오)
 - 사용자가 요구하는 정보를 원하는 시간에 볼 수 있도록 전송하는 멀티미디어 서비스
 - 프로그램을 시청하는 도중에 일시 정지, 느린 속도, 반복 재생 가능
 - TCP/IP(Transfer Control Protocol/Internet Protocol)
 - 인터넷에 연결된 기기종 컴퓨터들 간에 데이터를 주고 받을 수 있도록 하는 표준 프로토콜
 - 망의 일부가 파손되어도 남아 있는 망으로 통신이 가능한 신뢰성 있는 통신 규약

- 빅데이터 정의들

- 위키피디아 정의

- : 기존 데이터베이스 시스템의 데이터 수집-> 저장 -> 관리 -> 분석 기능을 넘어서는 정형·비정형 데이터 세트로부터 가치를 추출하고 결과를 분석하는 기술

- 국가전략위원회 정의

- : 대용량 데이터에서 가치 있는 정보를 추출하고 생성된 지식을 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술

- 삼성경제연구소 정의

- : 빅데이터란 기존의 관리 및 분석 체계로는 감당할 수 없을 정도의 거대한 데이터 집합으로 대규모 데이터와 관계된 기술 및 도구(수집·저장·검색·공유·분석·시각화 등)을 모두 포함하는 개념

→ (양적 + 질적) 데이터로서 의미를 지님

- 빅데이터 수집, 저장, 분석

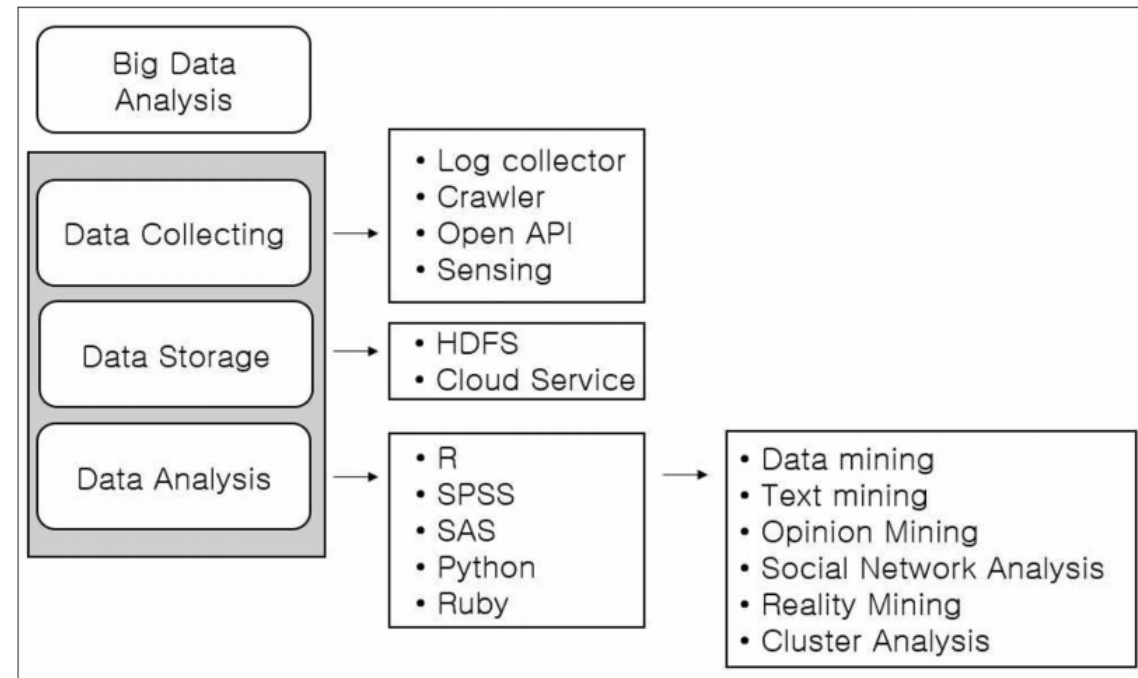
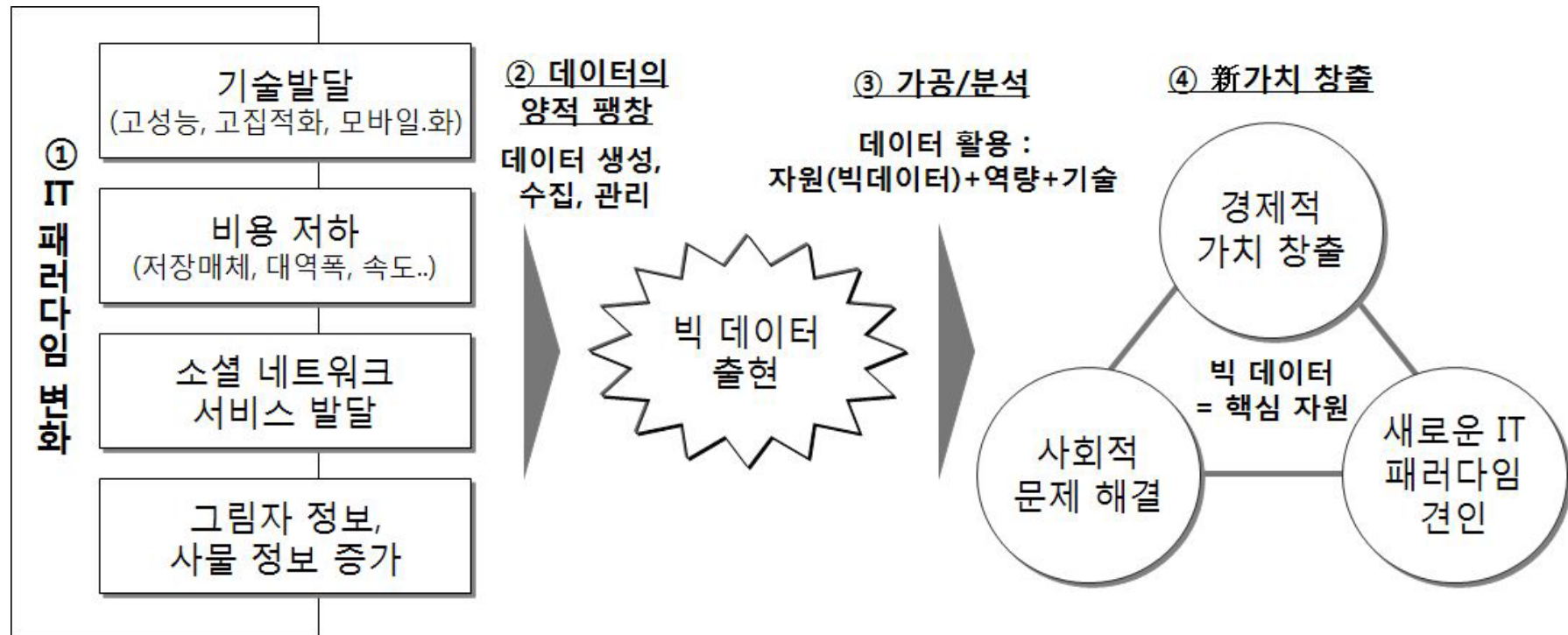


Fig. 1. Summary of big data analysis.

- 빅데이터 활성화 요인 3가지
 - 스마트폰 보급의 활성화 : 비정형 데이터 폭발적 증가
 - 클라우드 서비스 : 사용자 데이터 축적 -> 분석 -> 재활용 요구 증대
 - SNS 등의 소셜 미디어의 일상화 : 정보 유통 구조 새롭게 재편(쌍방교류)

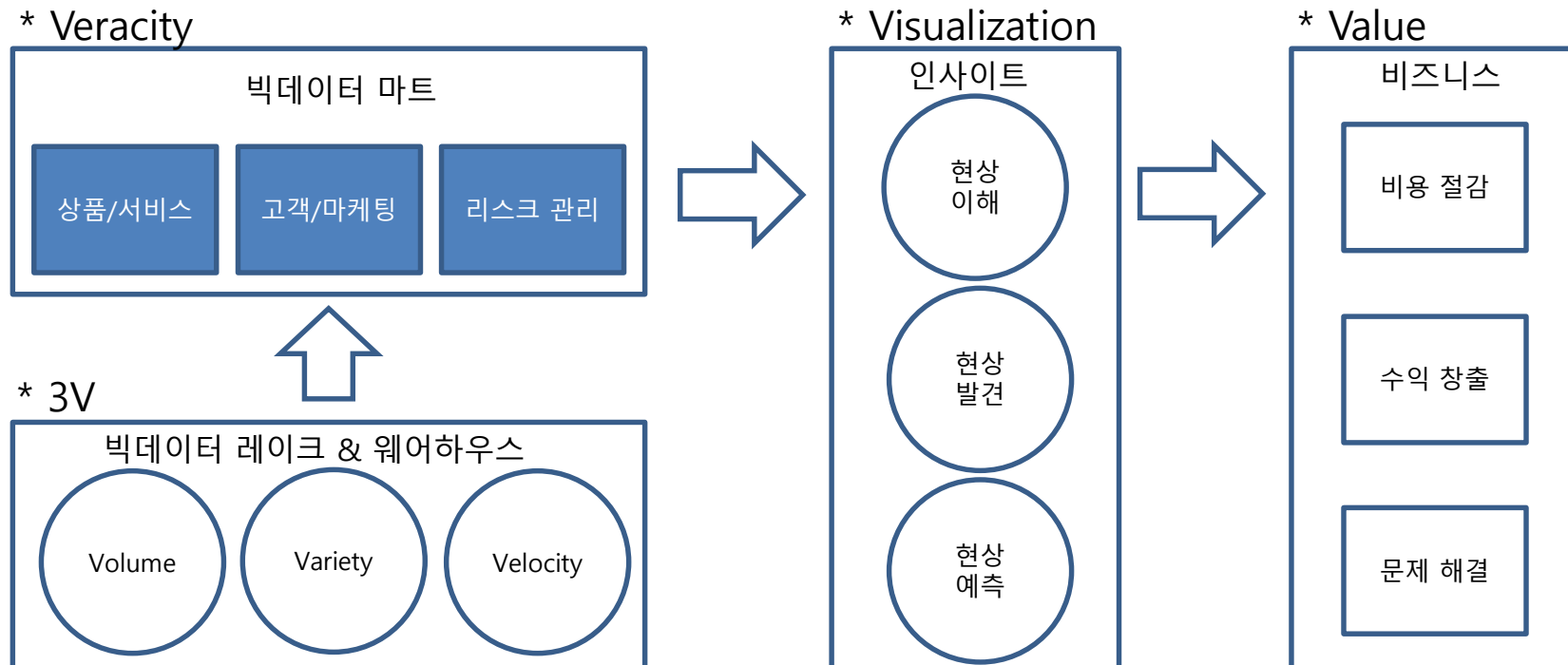


- 빅데이터 특성과 내용
 - 3V : 용량성(Volume), 다양성(Variety), 신속성(Velocity)
 - 5V : 3V + 정확성(Veracity, 신뢰정도), 가치성(Value)
 - 6V : 5V + 시각화(Visualization)
 - 7V : 5V + 신뢰성(Validity, 타당정도), 휘발성(Volatility)

구분	주요 내용
용량성(volume)	기하급수적으로 증가하는 디지털 정보량 -> 제타바이트(ZB, 1조1000억 기가바이트 (GB)) 시대
다양성(variety)	비정형화된 데이터의 증가 -> SNS, GPS, IoT, AI 데이터 등
신속성(velocity)	실시간성 정보 요구 및 증대 -> 사물 정보(센서, 모니터링), 스트리밍 정보, 모바일 정보의 실시간 처리
정확성(Veracity)	정밀한 정보 처리를 위한 빅데이터 분석 -> 무인화와 로봇 시스템의 질 높은 데이터 필요
가치성(Value)	창조와 가치 창출 -> 기업과 사회의 문제 해결을 위한 유용한 가치를 제공

- 빅데이터 목적 및 활용 방안[3]

- 빅데이터 목적 : 비용 절감, 수익 창출, 문제 해결
- 빅데이터 인사이트(통찰력) 3가지 유형 : 이해, 발견, 예측
 - 현상 이해 : 과거 발생한 일에 대한 이해와 원리 파악
 - 현상 발견 : 새로운 패턴 발견과 해석을 통해 무슨 일이 새롭게 일어났는지 파악
 - 현상 예측 : 이해와 발견을 기반으로 예측 모델을 만들어 미래에 발생할 현상 예측



- 성공적인 빅데이터 활용 전략
 - 빅데이터 성공적인 활용을 위한 3대 요소 - **자원, 기술, 인력**

<div>자원</div> <div>빅데이터</div>	<div>기술</div> <div>빅데이터 플랫폼</div>	<div>인력</div> <div>빅데이터 사이언티스트</div>
<ul style="list-style-type: none"> - 빅데이터 자원 확보 - 데이터 품질 관리 	<ul style="list-style-type: none"> - 데이터 저장 관리기술(NoSQL, ETL 등) - 대용량 데이터 처리(Hadoop, MapReduce 등) - 빅데이터 분석(자연어 처리, 의미 분석, 데이터 마이닝 등) 시각화 	<ul style="list-style-type: none"> - 수학, 공학 능력 - 경제학, 통계학, 심리학 등 다문화적 이해 - 비판적 시각과 커뮤니케이션 능력 - 스토리텔링 등 시각화 능력

참고 및 자료 출처

- [1] 윤경배 등, "4차 산업혁명의 이해 [2판]", 일진사, 2021
- [2] 한기준, 김기윤 등, "2020 시나공 정보처리산업기사 실기", 길벗, 2020
- [3] 김강원, "실무로 배우는 빅데이터 기술", 위키북스, 2020
- [4] TTA 한국정보통신기술협회 정보통신용어사전
<https://terms.tta.or.kr/main.do>