



빅데이터란 무엇인가



Index

01 빅데이터의 등장 배경

02 빅데이터란?

03 빅데이터의 예시 및 활용사례

04 빅데이터를 활용한 정책변화

1

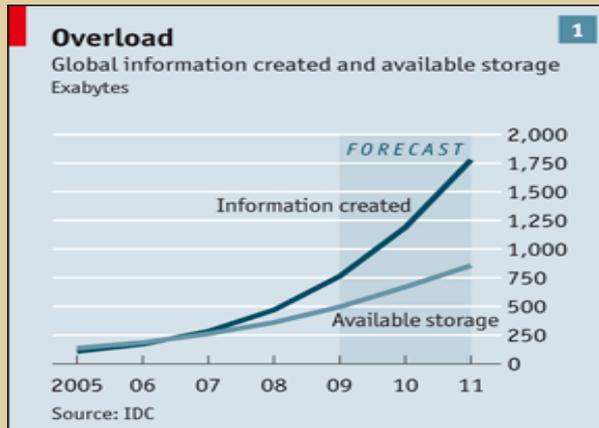
PART 01



빅데이터의 등장배경

01 빅데이터의 등장배경

- 2007, 생성된 디지털 정보량이 저장공간을 초과¹⁾
- 2010, 인터넷의 일상화와 함께, 데이터 홍수 현상에 직면²⁾
- 2011, Zeta byte 시대에 진입

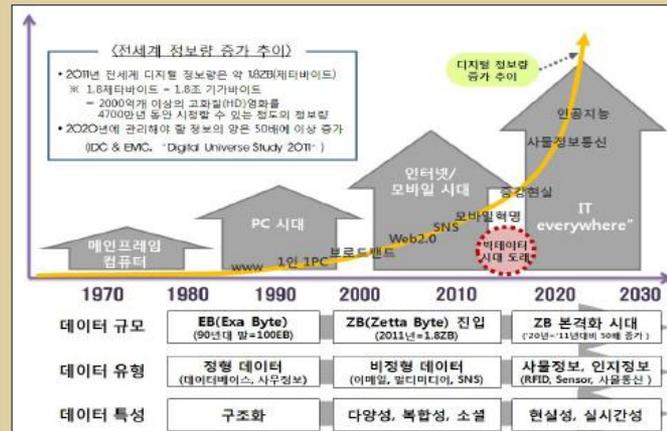
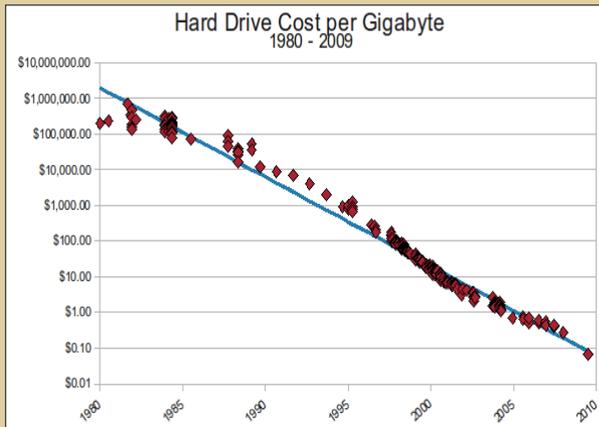


1) Economist Feb 25th 2010: 'Data, data everywhere'

2) 중앙일보 2011.6.30: '전 세계 데이터 빅뱅, 2년마다 2배씩 증가'

01 빅데이터의 등장배경

- 데이터 저장가격의 하락, 1980년 Gigabyte 당 약 10억 원
2010년 Gigabyte 당 약 100원³⁾
- 데이터 수집, 유지, 관리, 분석, 활용, 서비스 생태계 환경 촉진⁴⁾



3) <http://www.mkomo.com/cost-per-gigabyte>

4) 한국정보화진흥원 (2012) IT&Future Strategy 18호 - 빅데이터의 새로운 가능성과 대응 전략

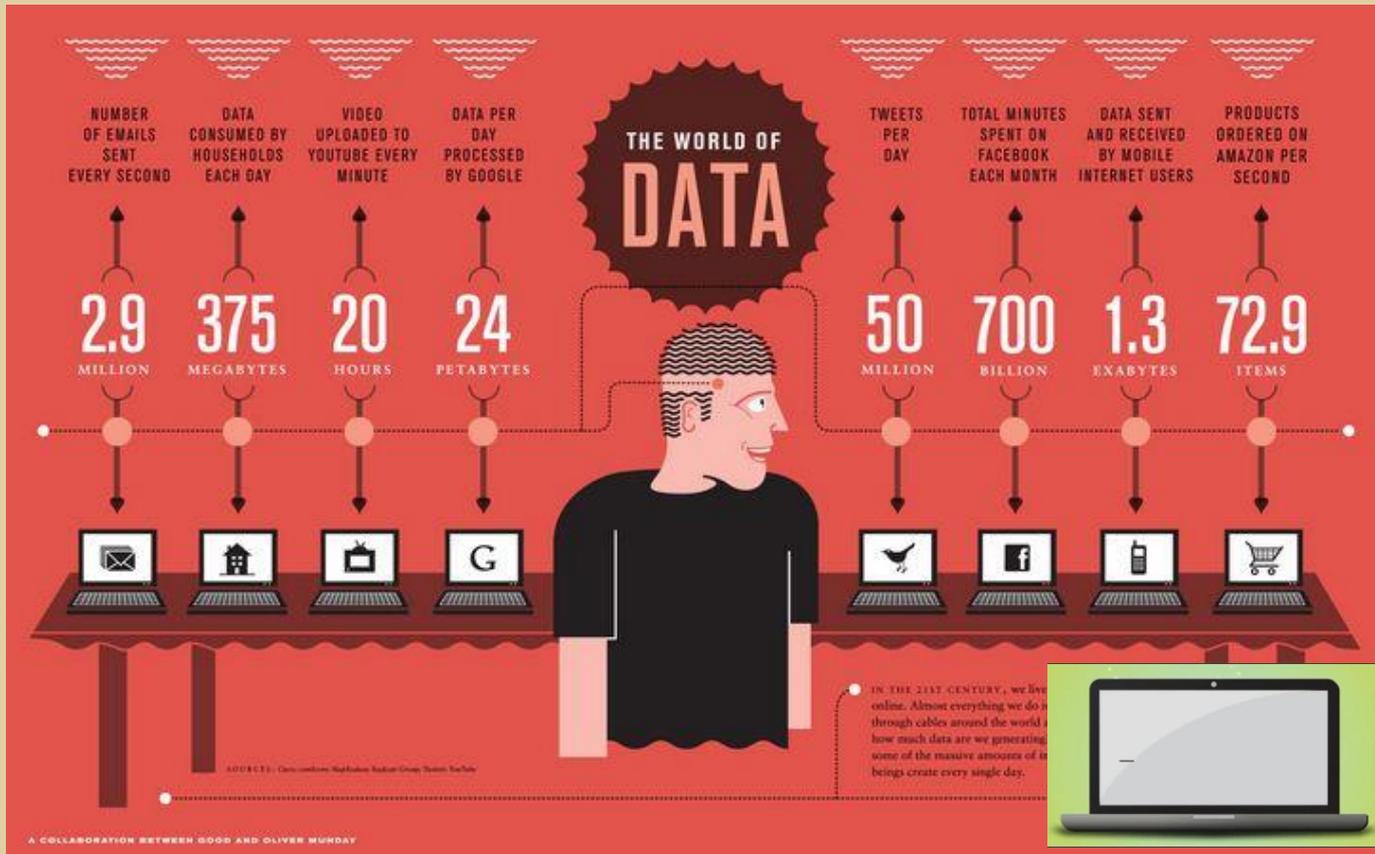


PART 02



빅데이터란?

빅데이터란?



02 빅데이터란?

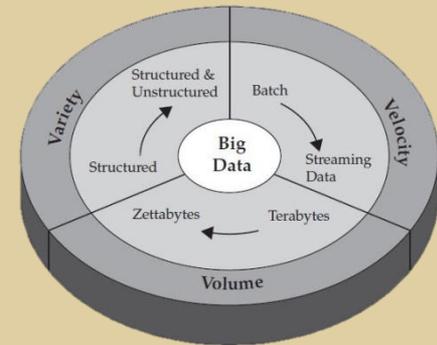
- 데이터 형식 및 입출력 속도의 변화로 인해 기존의 분석 방법으로는 수집, 저장, 검색, 분석하기 어려운 데이터. 규모가 방대하고, 종류가 다양하며, 데이터의 생성-유통-이용이 실시간으로 이루어짐

- 빅데이터의 주요 요소 '3V' 5)

Volume: 수집되고 분석되는 데이터의 양

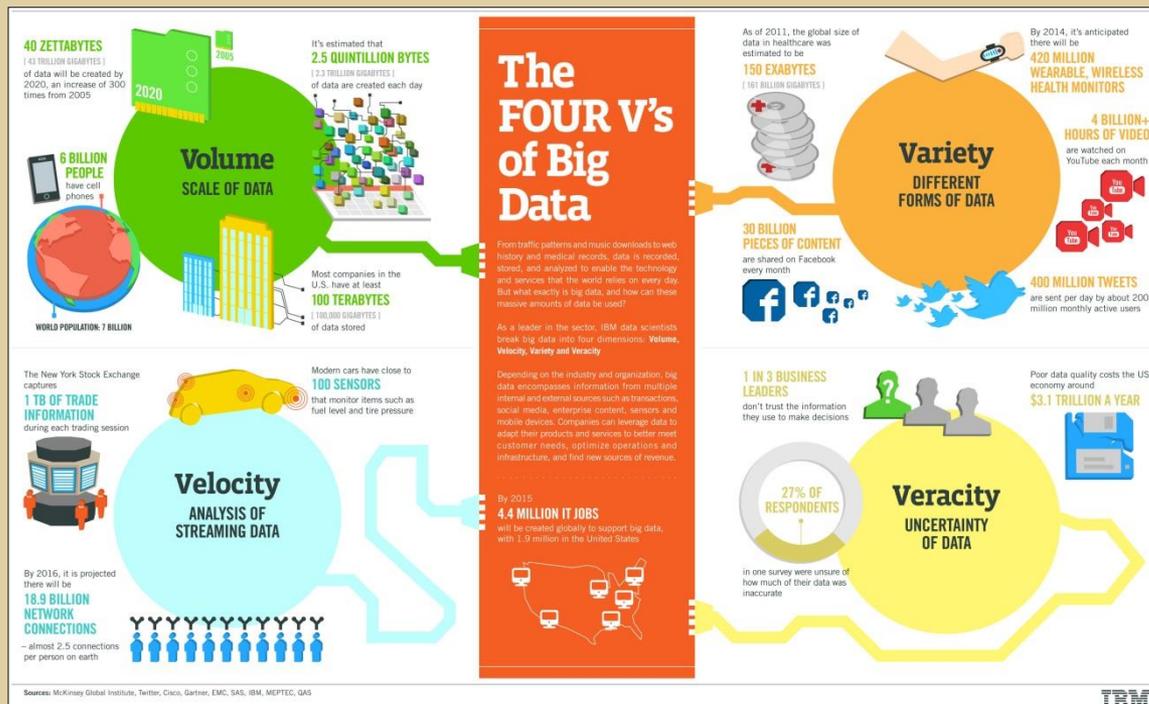
Variety: 정형/비정형 데이터

Velocity: 데이터가 처리되는 속도



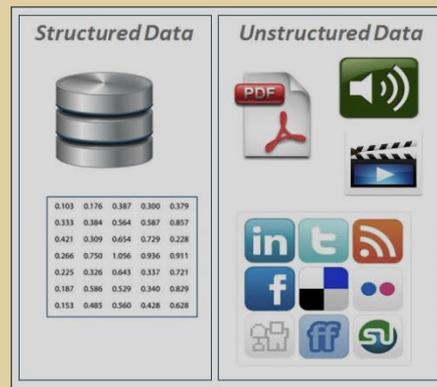
빅데이터란?

■ '3V' + Value, Veracity, Complexity...6)



02 빅데이터란?

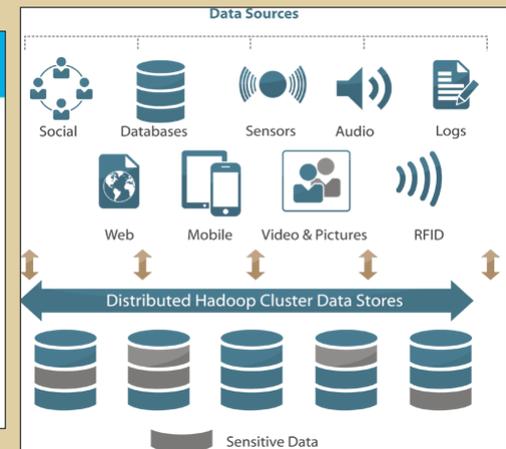
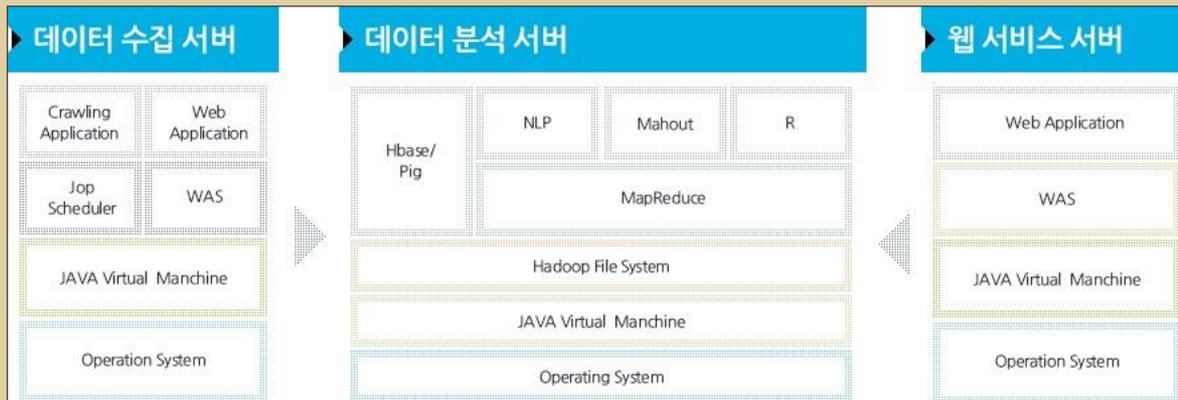
- 인터넷의 발전, 모바일 기기의 진화, 소셜 네트워크의 출현으로, 예전에는 수집되지 않았던 비정형 데이터가 자동으로 발생
(웹 이용패턴, 교통정보, RFID, CCTV, 로그기록, 동영상 등)



http://www.denodo.com/en/product/components_features.php#

02 빅데이터란?

- 데이터 수집, 저장, 처리, 분석 기술이 함께 발전하고 있음



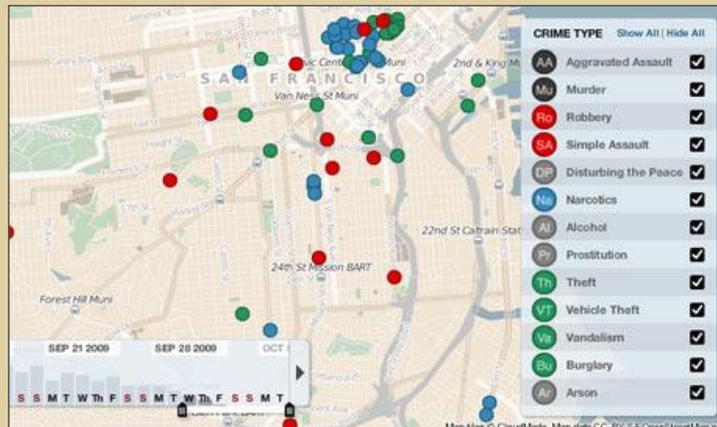
A large, stylized number '3' in a light yellow color. The top curve of the '3' is open on the left side, and the bottom curve is open on the right side. The text 'PART 03' is centered within the middle of the '3'.

PART 03

빅데이터의 예시 및 활용사례

빅데이터의 예시 및 활용사례

- 샌프란시스코 범죄 정보 활용 사례: 경찰의 범죄 지도와 과거 범죄 발생 지형 및 유형을 분석, 범죄 예측 정확도 71% 수준⁷⁾



7) <http://sf-police.org/index.aspx?page=1618>

03

빅데이터의 예시 및 활용사례

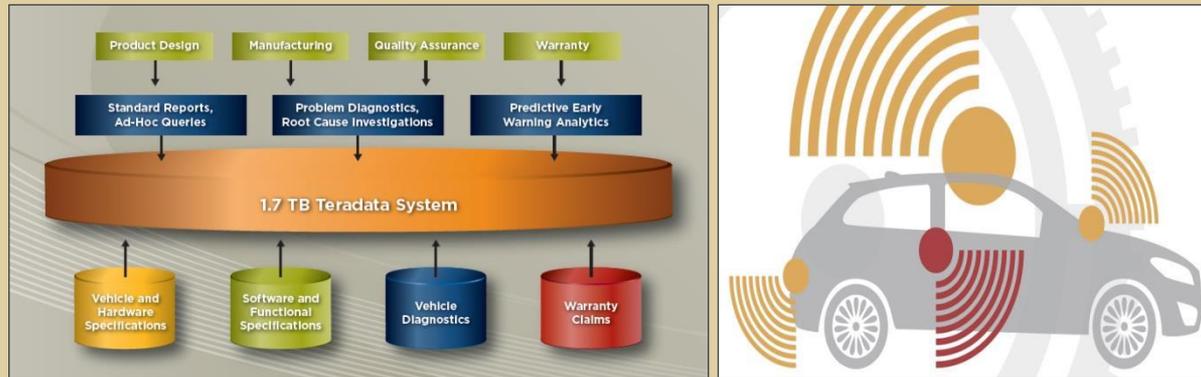
- 온타리오 병원, 미숙아 인큐베이터 모니터링 장비에서 미숙아들의 혈압, 체온, 심전도 등 생리학 데이터를 수집, 분석하여 실제 의료진보다 6~24시간 먼저 신생아의 감염사실을 알아내고 처방⁸⁾



8) <http://www.cnn.com/id/101032950>

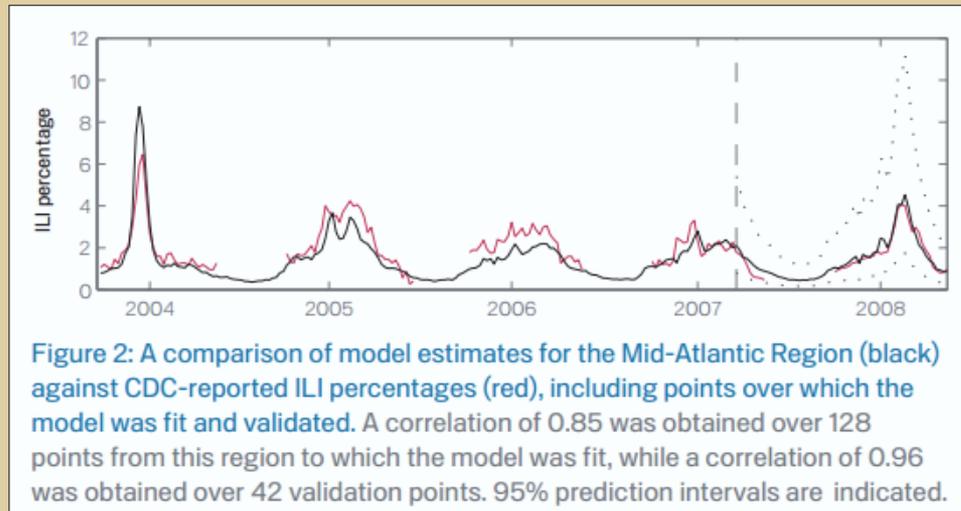
빅데이터의 예시 및 활용사례

- 자동차사 Volvo, 차 내부의 RFID센서를 통해 부품의 상태 및 안정도 등 다양한 차량 정보를 수집하고, 운전자 요구사항, CRM, 공장데이터와 통합 분석하여 차량 결함을 파악, 대응⁹⁾



빅데이터의 예시 및 활용사례

- Google, 독감과 관련된 주제를 검색하는 사람의 수와 실제 독감 증상이 있는 사람 수 간에 밀접한 관계가 있으며, 연도별 유행 수준과 비교할 경우 정확도가 유의미한 수준¹⁰⁾



10) Ginsberg, Jeremy, et al. "Detecting influenza epidemics using search engine query data." *Nature* 457.7232 (2008): 1012–1014.

A large, stylized number '4' composed of thick, gold-colored geometric shapes. The top part is a triangle pointing right, followed by a vertical bar, and the bottom part is a horizontal bar with a vertical stem extending downwards.

PART 04

빅데이터를 활용한 정책 변화

04

빅데이터를 활용한 정책 변화

- 공공데이터 활용과 Open Government 확산
- 선진국에서는 이미 활발하게 빅데이터를 활용한 정책 수립을 시행중
예: 영국 ‘Open Knowledge Foundation’ , 영국 정부가 개방한 공공데이터로 개인의 수입과 세금 납부액 및 세금 사용 분야에 대한 정보 제공¹¹⁾



11) 한국정보화진흥원 (2012) IT&Future Strategy 6호 - 빅데이터 시대 공공분야 적용방향

04

빅데이터를 활용한 정책 변화

- 서울시 심야버스: 유동인구 행동패턴 및 교통카드, 이동통신사의 통화량 분석을 통해 심야버스 노선 정책에 반영¹²⁾



12) 시사인 2013.11.18 '어쩐지 심야버스가 척척 오더라' <http://www.sisainlive.com/news/articleView.html?idxno=18411>

04

빅데이터를 활용한 정책 변화

- 정부3.0, 빅데이터 기반 과학적 행정 서비스 제공
경찰청 범죄예측 지리적 프로파일링 시스템

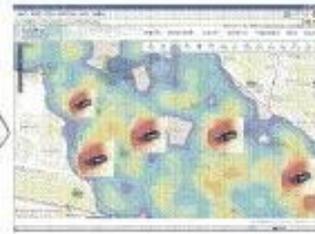
· 범죄위험지역, CCTV위치, 범죄발생 이력데이터 등을 종합적으로 분석하여 순찰차 이동경로 조정 및 범죄예방전략 수립



GEOPROS
범죄위험지역 예측



순찰차
신속배치 시스템(IDS)



· 순찰차 이동경로와 범죄 발생지도 비교
· 범죄위험지역에 순찰차 배치전략 수립

Thank You