



데이터와 금융혁신

2019. 10. 2.

연구위원 권민경



1 데이터의 개념

2 데이터의 활용

3 데이터와 혁신

4 시사점

데이터의 개념

○ 데이터 경제

- › 데이터는 4차 산업혁명 시대 중요한 생산요소
 - Digitalization, 빅데이터, 인공지능, 사물인터넷, 클라우드 등의 토대
 - '21세기 원유', '자본과 노동 등 기존 생산요소를 능가하는 경쟁원천' 등으로 평가
- › 2018년 8월 '데이터 경제로의 전환' 선언
 - 정부는 데이터·AI 경제를 혁신성장 정책 중 3대 전략투자 분야로 선정

○ 데이터의 개념

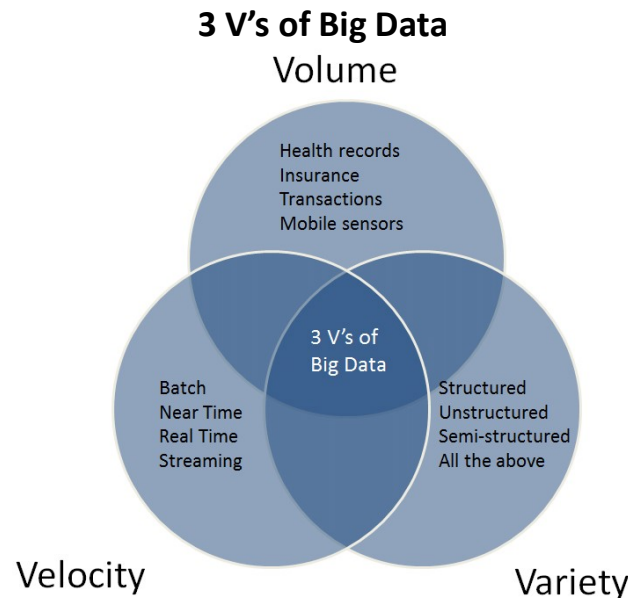
- › 데이터는 정보 생성을 위해 필요한 것
 - 데이터(data) → 정보(information) → 지식(knowledge)
 - (예) 과거 100년 간 세계 곳곳의 온도 측정(데이터) → 지구 온난화 진행 중(정보)
- › 컴퓨터로 처리할 수 있는 형태로 저장된 자료
 - 데이터의 수집 및 분석 방법의 변화
 - (1) 사람이 수기로 데이터 축적 → 수기 분석
 - (2) 사람이 컴퓨터에 데이터 입력 → 컴퓨터로 분석
 - (3) 컴퓨터, 기계, 센서 등에서 데이터 축적 → 컴퓨터로 분석

데이터의 개념

● 관련 핵심어

› 빅데이터(big data)

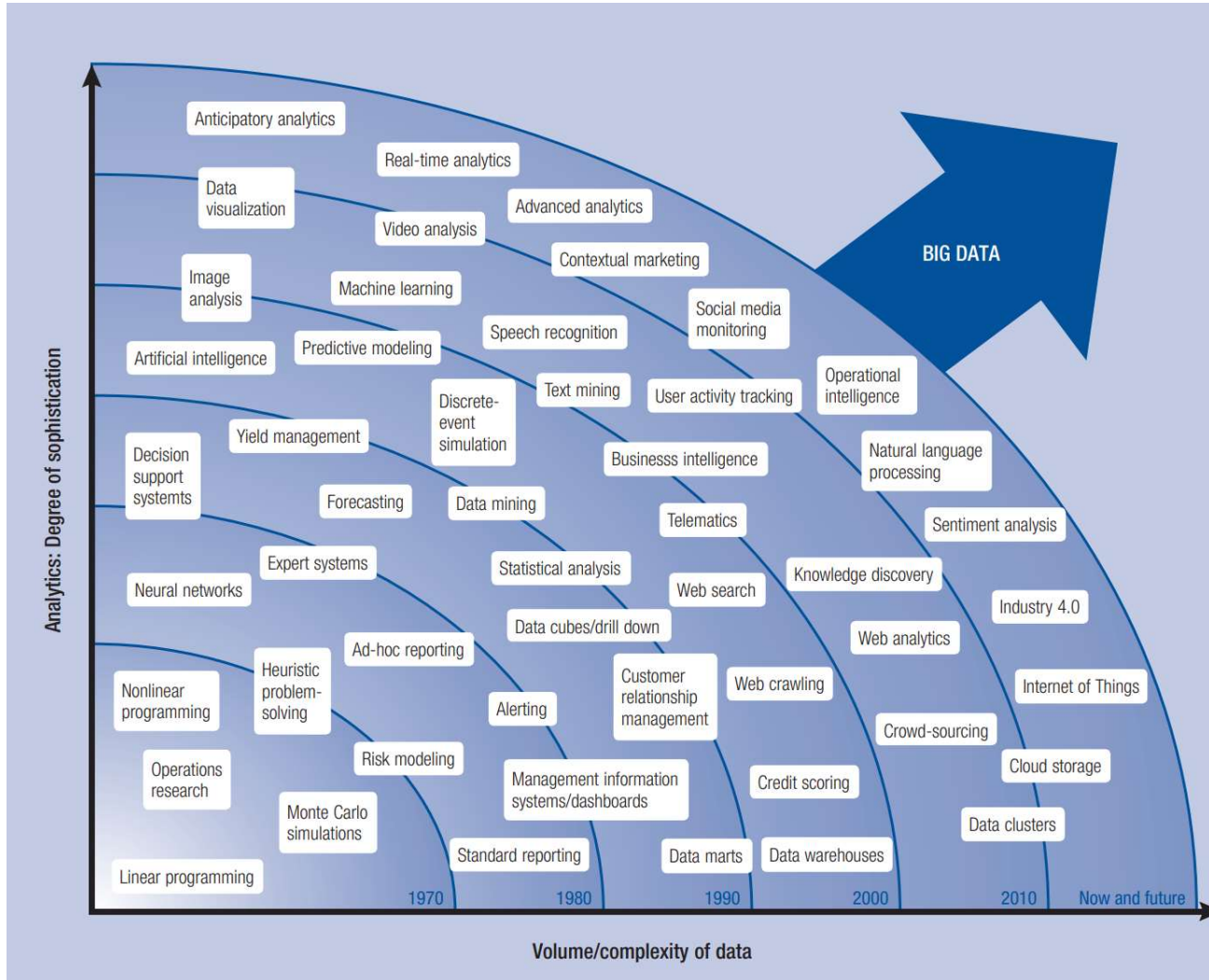
- 전통적 방식으로 저장하거나 처리하기에 지나치게 큰 데이터의 집합(dataset)
- 크기(Volume), 속도(Velocity), 다양성(Variety) 등 3V로 그 특징을 설명
- 센서, IoT 장치, 스마트폰, 무선네트워크, 로그(log), 카메라, 마이크, RFID, SNS 플랫폼 등의 발달로 데이터가 폭발적 증가함에 기인
- '머신러닝'과 같이 빅데이터를 분석하는 알고리즘도 빠르게 발달



자료: Austin and Kusumoto(2016) (Laney(2001)를 도식화한 표임)

데이터의 개념

데이터 기반 분석의 진화



데이터의 개념

- 구성요소: 수집 → 저장 → 처리 → 분석 → 표현
 - 빅데이터는 분산저장, 분산처리 등 차별화된 방식 사용
 - 실시간처리 방식이 점차 늘어나는 추세
 - 전통적 분석이 '인과관계'에 집중한 반면, 빅데이터에서는 '연관·상관관계' 위주 분석

빅데이터 플랫폼의 구성요소별 주요 기능

수집	저장	처리	분석	표현
비정형 수집 정형 수집 ETL Web Robot EAI, ESB, FTP 등 Open API	원본 데이터 NoSQL 메모리 검색엔진 데이터보안	배치처리 실시간처리(CEP)	텍스트분석 Machine Learning 통계 Data Mining SNS 분석 예측분석(알고리즘)	시각화

자료: 안춘모(2017)

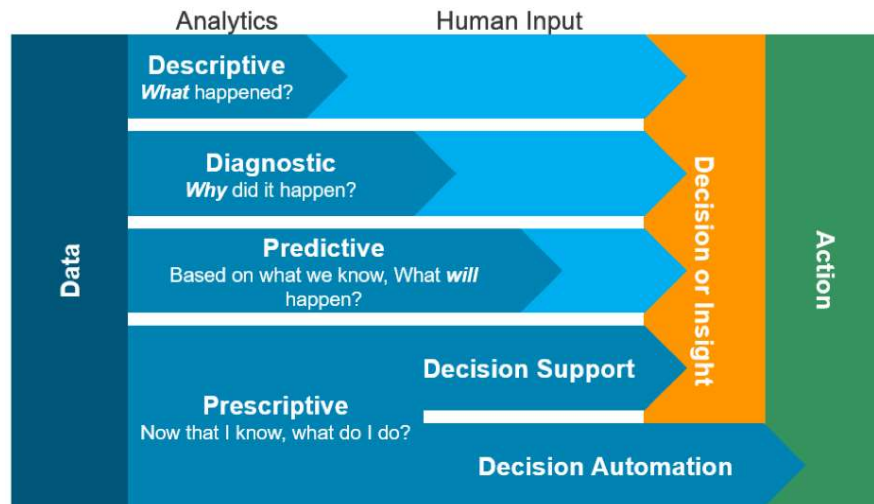
› 클라우드(cloud)

- 데이터의 저장 및 컴퓨팅 파워 등을 온디맨드(on-demand) 형식으로 외부에서 아웃소싱해 사용하는 서비스
- 빅데이터 활용에 있어서 핵심 인프라 역할을 수행

데이터의 개념

데이터의 활용

- › 문제를 파악하고 이를 해결하기 위해 데이터를 활용
- › 각각의 단계에서 데이터의 활용 가능성 증대
 - (descriptive) 어떤 문제가 있는가?
 - (diagnostic) 문제가 왜 발생하였는가?
 - (predictive) 향후 무슨 일이 일어날까?
 - (prescriptive) 어떤 행동을 취해야 하는가?
- › '직관'이 아닌 '데이터'에 기반한 의사결정으로 진화

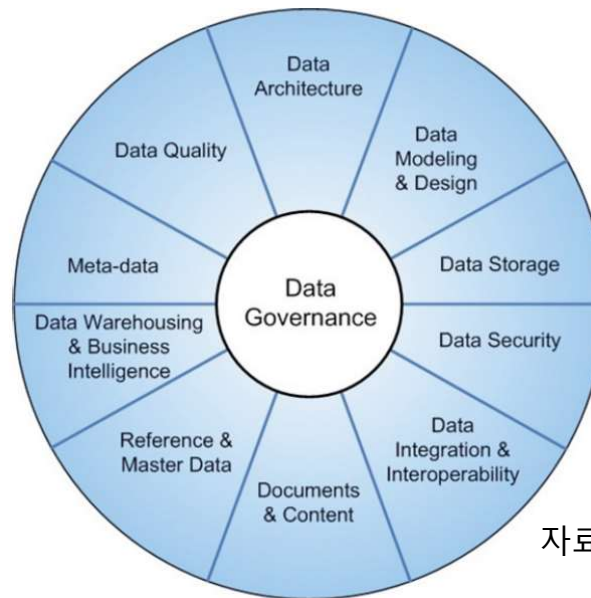


자료: Gartner(2015)

데이터의 개념

데이터 관리

- 고품질 데이터 확보, 적극적 활용을 통한 가치 창출을 목적으로 함

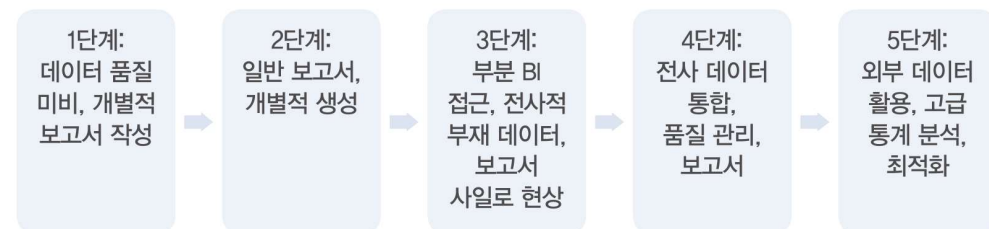


자료: 조완섭(2017)

데이터의 공유 방식의 진화

- (1) 개별 존재
- (2) 부서별 산재(silo)
- (3) 전사 공유
- (4) 외부 데이터 결합

기업의 분석경쟁력 5단계



자료: 김옥기(2018)



1 데이터의 개념

2 데이터의 활용

3 데이터와 혁신

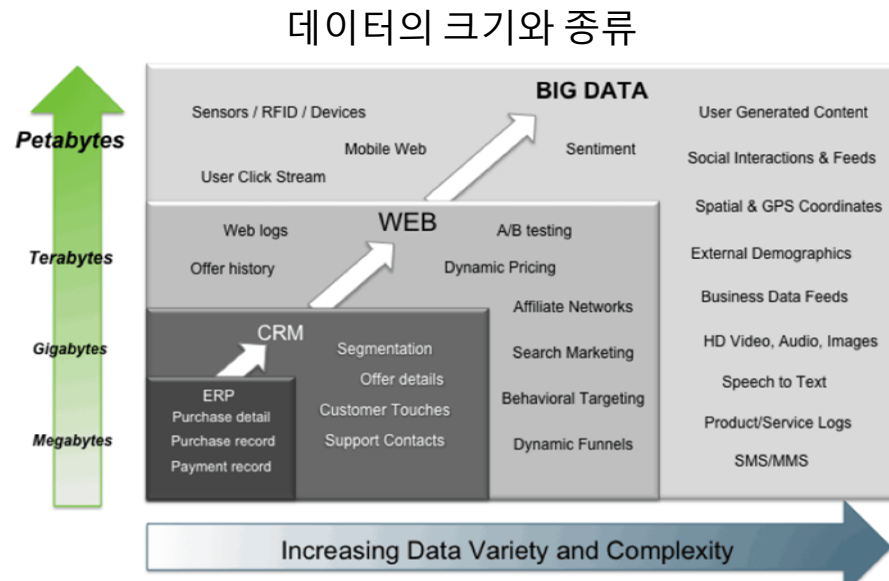
4 시사점

데이터의 활용

○ 일반 기업 및 공공기관에서의 활용

› 데이터의 종류

- (1) Enterprise Resource Planning
- (2) Customer Relationship Management
- (3) 웹 로그
- (4) 센서, RFID, 모바일, 사용자 클릭 스트림, SNS 등 각종 플랫폼 생산 데이터 등



자료: Teradata

데이터의 활용

› 데이터의 활용 목적

- (일반) 매출·수익 등 성과 분석
- (제조) 주문·재고·품질 관리, 유지보수
 - GE는 매출액의 75%가 자사 제품에 부착한 센서데이터를 통한 유지보수에서 발생
 - 지멘스는 설비에서 발생한 데이터를 분석하고 생산라인을 재조정하여 생산량 8배 증가
 - 볼보는 자동차에 탑재된 센서 데이터를 통해 차량 관리의 효율성 제고
- (판매) 매출 패턴 및 고객 분석, 마케팅
 - 신규 고객 유치, 고객 이탈 방지, 추가/연계 판매(up-sell, cross-sell)
 - 콜센터에서 고객의 과거 접촉 이력을 활용하여 성향을 분류하고 맞춤형 응대, 피드백 결과 수집하여 차세대 행동 통계 모델로 발전
 - 웹로그 및 소셜플랫폼 분석 결과를 반영하여 신상품 개발
- (공공) 의료, 통신, 복지, 교통, 환경, 범죄·사기 감지
 - 범죄가 일어날 시간 및 장소 예측, 범죄 용의자 추정 등
 - 교통예측, 유료도로 통행료 측정, 실시간 교통정보 제공 서비스
 - 조류독감 확산 조기 대응

데이터의 활용

○ 금융기관에서의 활용

- › 공장에서 기계 내장 센서들이 데이터를 축적하듯,
금융에서도 시장 참여자와 인프라 시스템들이 수많은 데이터를 생성
- › 빅데이터 활용 배경 (Gutierrez, 2014)
 - 금융상품을 거래하거나 서비스를 제공받는 채널이 온라인으로 변화
 - 오프라인에서 축적되지 않았던 고객 관련 데이터가 급증
 - 온라인 거래의 편의성으로 인해 금융상품 거래 빈도가 크게 늘어남
 - SNS와 같이 새로운 플랫폼 데이터가 생성
 - 소속 집단 및 인간 관계를 통해 정보 유추 가능
 - 금융기관에 대한 위험한도 규제 및 보고의무 강화
- › 데이터의 종류
 - 시장 관련
 - 시장 주문 및 거래내역, 보도자료, 공시자료, 애널리스트 보고서, 뉴스, 소셜미디어
 - 각종 위성 및 센서, 웹 스크레이핑 등으로 수집한 자료, 외부 데이터베이스
 - 고객 관련
 - ATM, 콜센터, 지점, 온라인 이용/방문 내역
 - 모기지, 신용카드 등 금융상품 거래내역

데이터의 활용

› 활용목적

- (제조) 시장 위험 및 심리(sentiment) 파악, 신용 위험 분석 및 가격 결정
 - 거래상대방에 대한 신용위험 측정
 - 모기지 등 복잡한 금융상품의 세부 위험 요인을 파악하고 가격 결정
- (판매) 소비행태 및 위험성향에 따른 고객 분류, 맞춤형 서비스 제공
 - SNS 통해 고객의 약혼 사실 파악 → 신혼부부 대출 상품 제안
 - 여의도에서 직장을 다니는 30대 여성 → 동일 그룹의 최선호 금융상품 제안
- (운영) 운영비용 절감, 강화된 금융규제 및 보고의무 준수, 재무 유동성 관리, 내부통제 강화, 이상거래 및 사고 탐지
 - 10분 전 서울에서 ATM 입금한 고객 계좌, 부산에서 인출 시도 → 이상신호 발생
 - 웹로그 분석을 통해 고객이 금융서비스 받는데 소요되는 시간 단축



1 데이터의 개념

2 데이터의 활용

3 데이터와 혁신

4 시사점

데이터와 혁신

○ 오픈데이터

- › 누구나 접근 가능하고 활용 및 공유할 수 있는 데이터(Open Data Institute)
- › 공공기관 데이터는 기본적으로 누구나 제약 없이 활용할 수 있어야 함
 - '13년 G8 국가 정상회의에서 오픈데이터 헌장(Open Data Charter) 마련
 - 원칙적 개방(open data by default)
 - 품질 및 규모 측면에서 우수(quality and quantity)
 - 누구나 활용 가능(usable by all)
 - 거버넌스 향상(releasing data for improved governance)
 - 혁신 지원(releasing data for innovation)
 - 법적 개방성과 기술적 개방성을 요구(Open Knowledge Foundation 참조)
 - 국가안전보장 등 공익이나 타인의 권리를 침해하는 경우는 적용 제외

데이터와 혁신

› 공공기관 데이터의 개방

• 미국

- 2009년 Transparency & Open Government 정책의 일환으로 오픈데이터 도입
 - 연방정부에서 공공데이터 포털 data.gov을 운영할 뿐 아니라 주정부에서도 적극 참여
 - Zillow, WeatLaw 등 민간기업이 공공데이터를 활용해 수익을 창출*
- * 박경현 외(2017) 참조

• 영국

- 2010년 데이비드 캐머론 총리 내각 출범과 함께 오픈데이터 추진
 - 공공데이터 포털 data.gov.uk을 운영
 - Open Data Strategy('14)를 통해 공공데이터 활성화 노력
 - Open Corporates, Spend Network 등 민간기업이 이를 활용해 수익 창출*
- * 정보통신산업진흥원(2014) 참조

데이터와 혁신

- 한국

- 공공데이터법 제3조 1항 "공공기관은 누구든지 공공데이터를 편리하게 이용할 수 있도록 노력하여야 하며, 이용권의 보편적 확대를 위하여 필요한 조치를 취하여야 한다."
- 정부의 「데이터 산업 활성화 전략」('18.6)
 - 민간 수요가 높은 데이터를 국가중점데이터로 선정, 조기 개방 확대
 - 공공재 성격의 민간데이터 개방 확대:
통신, 포털 검색, 뉴스, 유통, 금융거래 등 산업적 활용도 높은 민간 빅데이터 개방
- 공공데이터포털(data.go.kr), 서울열린데이터 광장 등 구축

주요 분야 시범사업(안)

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	

자료: 정부, 「데이터 산업 활성화 전략」('18.6) (제시된 표에서 일부 발췌)

데이터와 혁신

- 주요 공공데이터(한국정보화진흥원)
 - (문화/관광) 전국 도서관 대출 데이터, 유물 3D 스캔 원시데이터, 관광 및 상권 정보
 - (교통/물류) 주차장, 도로이정표 및 교통정보, 전기차 충전소 위치
 - (환경/기상) 기상정보 서비스
 - (보건/의료) 병원 및 약국, 응급의료기관 정보, 환자데이터셋
 - (산업/고용) 전자공시, 국민연금 정보
 - (식품/건강) 화장품 특성성분, 식품영양, 사료배합 정보
 - (교육) 학원, 공공교육 데이터, 학교 일정 및 급식 정보
 - (국토관리) 건축물대장, 토지대장, 실거래가, 공시가격, 인구분포
 - (농축수산) 도매시장 경락가격
 - (사회복지) 장애인 편의시설/활동지원기관, 봉사참여 및 배움 정보
- 서울 열린데이터광장 인기 데이터
 - 구/동별 주민등록인구
 - 지하철/버스 노선/정류장별 승하차 인원 정보
 - 지하철 실시간 도착/위치 정보
 - 권역별 실시간 대기환경, (초)미세먼지 예경보 현황
 - 유동인구 및 사업체 정보

데이터와 혁신

○ 민간데이터의 거래

- › 미국·중국에서 데이터거래소 또는 개별 접촉을 통한 민간 데이터의 거래 활발
 - 이들 국가는 개인정보의 활용이 비교적 용이
 - 액시엄(Acxiom)은 전 세계 25억명에 대해 총 1만개에 달하는 특성 데이터를 보유
 - 금융·유통회사 등에서 내부 데이터와 결합, micro targeting 등 주로 마케팅에 활용
- › 한국은 상대적으로 개인정보 규제가 엄격하여 민간 데이터 거래가 제한적
 - 데이터스토어 인기 데이터(한국데이터산업진흥원)
 - 다이닝코드 > 전국 음식점 운영시간 및 메뉴, 가격
 - BC카드 > 전국지역/업종별 소비데이터
 - KB카드 > 고객 프로파일 그룹별 매출 데이터
 - 모바일티머니 > 대중교통 승하차 정보
 - SK텔레콤 > 서울 요일별 유동인구 데이터

데이터와 혁신

○ 마이데이터

› 배경

- 정보주체인 개인의 활용도는 낮은 상황
- 정보 비대칭으로 인해 비경쟁적 시장 환경 조성
- 현행 계좌통합조회 서비스에서 데이터 전송 방식은 불완전

› 정책 도입

- 개인에게 '(금융)기관→제3자'로의 데이터 전송 요구권(right to data portability) 보장
- 의료, 금융, 통신 등 분야에서 정보주체가 자기정보결정권을 행사
 - 금융기관의 API 개발 및 공개 의무화

주요 분야 시범사업(안)

분 야	내 용	참여
[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)

자료: 정부, 「데이터 산업 활성화 전략」(’18.6)

데이터와 혁신

› EU

- Right to data portability(데이터 전송 요구권) 도입
 - 본인 개인정보를 제3자에게 활용도 높은 형식*으로 제공할 것을 요청할 수 있음
*structured, commonly used, machine-readable and interoperable format
- API를 통한 제3자의 읽기, 쓰기 기능을 허용
 - 읽기: 잔액과 거래내역 등 금융기관이 축적한 고객 데이터를 제3자가 조회
 - 쓰기: 고객을 대신하여 제3자가 금융기관에 지급(payment)을 지시
- AIS와 PIS*를 법적 지급 서비스(payment service)로 편입
* Account Information Service, Payment Initiation Service

› 영국

- EU의 읽기, 쓰기 기능을 포함 뿐 아니라 금융상품 정보까지 API 공개 의무화
- 결제계좌 상품에 대해 가격비교사이트 방식의 비교 서비스 활성화

› 호주

- 읽기, 금융상품 정보의 공개 의무화
- 공개 대상 금융상품 범위 넓음

데이터와 혁신

› 한국

- 마이데이터 정책('18.7)
 - 읽기 기능 포함
 - 개인신용정보이동권(right to data portability) 도입
 - '금융회사→제3자' 데이터 전송 시 표준API 활용
 - 본인신용정보관리업 신설
 - 통합조회서비스, 정보계좌 업무, 데이터 분석/컨설팅, 투자자문/일임, 금융상품 자문 등
 - 공개 대상 금융상품 범위(안)
 - (은행·상호금융·저축은행·보험사) 예금계좌/카드 입출금 내역, 대출/보험계약 정보
 - (증권사) 투자자예탁금·CMA 등 입출금 내역, 금융투자상품(주식·펀드·ELS 등)의 총액
 - (전기통신사업자) 통신료 납부내역
- 한국의 공동 결제시스템(오픈뱅킹) 정책('19.2)
 - 쓰기 기능 포함

데이터와 혁신

○ 데이터 결합 및 활용 사례

- › 부동산 매매 관련 종합정보 서비스(예: Zillow)
 - 실거래가, 면적 및 구조, 관리비 등의 공공데이터를 수집
 - 매매호가 등 내부 데이터와 결합하여 부동산 수요자에게 폭넓은 서비스 제공
- › 심야버스 운행 노선 최적화
 - 통신회사의 심야 통화 장소와 청구지 주소 데이터를 통해 경로 추정
 - 야간 택시 승하차 데이터 활용
- › 통합 대기환경 지수 제작
 - 다양한 출처에서 대기환경 데이터 수집 → 차원 축소 → 지수화
- › 질병 극복(Patients Like Me 사례)
 - 동일한 질병을 가지고 있는 환자들이 경험을 공유할 수 있는 소셜 네트워크
 - 정형화된 질병 관련 데이터를 축적하여 질병 극복 연구에 활용
 - 현재 60만명의 가입자, 2,800여 개의 질병상태, 4천3백만 건의 질병데이터가 등록
 - 환자들이 익명으로 올린 데이터를 제약사 등에 판매

데이터와 혁신

○ 금융 산업에서의 혁신

› 금융상품 제조 관련

- 머신러닝으로 수많은 시장 데이터를 학습시켜 자산가격결정에 활용(Gu et al., 2018)
- 이미지 인식과 자연어 처리, SNS 플랫폼 등 기존에 활용하지 않던 새로운 데이터를 수집하여 활용
- 헤지펀드가 제공하는 데이터와 플랫폼을 통해 자산가격을 예측하려는 경쟁 시도
- 정교한 위험 측정 통해 시장 내 자본 배분 효율성 높아질 것으로 기대

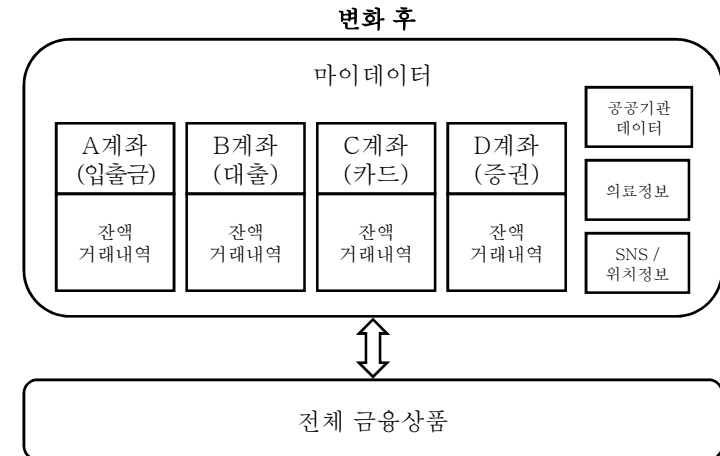
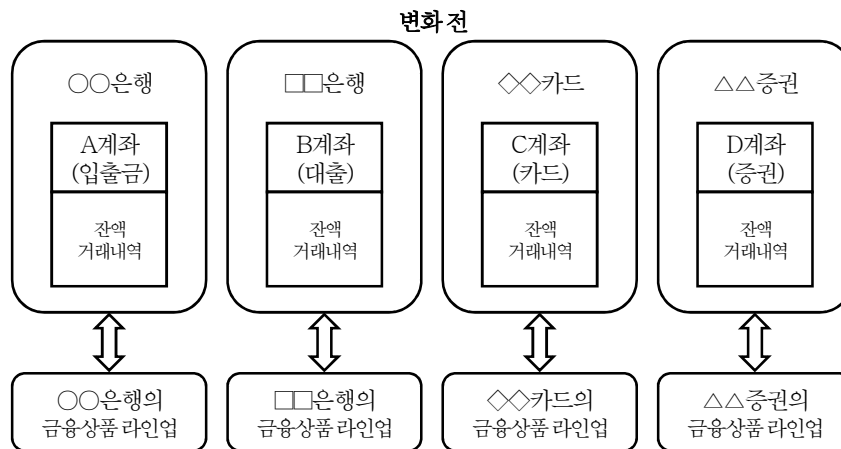
› 신용평가 방식의 개선

- 기존 방식의 문제점
 - 기존 정량적 분석에 쓰이는 데이터는 금융 거래 및 공과금 납부 및 연체 내역 등으로 한정
 - 해당 거래가 많지 않은 대출자의 경우 실제 상환능력과 무관하게 낮은 신용등급 부여됨
- 빅데이터를 활용하여 실제 상환능력을 확인
 - 전통적으로 활용하는 수십 가지 정량 지표 외 수천 개의 빅데이터를 머신러닝으로 분석
 - SNS, 이메일, 문자메시지 내역 등을 통해 대출자의 감정 상태, 행태, 사회적 관계 등 반영
 - 기존 저신용자 중 상당수가 충분한 상환능력을 입증하여 대출을 받음

데이터와 혁신

› 소비자의 합리적 의사결정 지원

- 본인 데이터의 통합조회는 효율적 자산배분을 유도
 - 리볼빙/연체/초과인출(overdraft) 등 감소
- 소비패턴이나 재무현황, 위험성향 등 파악하여 맞춤형 재무 컨설팅 제공
 - 고객 데이터 대입하여 최적 조건의 금융상품 추출
 - 이용 중인 상품에 대해 동일한 혜택을 제공하지만 비용이 저렴한 대체상품 추천
 - 동일 연령대, 비슷한 수입 규모의 소비자들이 선호하는 인기상품 제시



데이터와 혁신

› 고객 경험 제고

- 실시간 서비스 확대
 - 데이터가 API를 통해 실시간으로 전송
- 통신·의료·SNS 데이터 등을 결합하여 생애주기 및 위치기반 서비스 제공
 - 고객 상황에 적합한 맞춤형 상품의 개발 및 판매
 - 금융과 타 산업 데이터 간 결합은 사회적으로 큰 부가가치 창출 가능
- 다른 산업과의 연동하여 매끄러운(seamless) 고객경험 창출
 - 유통, 항공, 숙박 등 산업에서 금융서비스는 고객 경험에 큰 영향을 미침

› 금융당국에서의 활용

- 어떠한 정책과제가 가장 시급한지 우선순위를 정하는데 사용
 - 미국 금융소비자보호국은 SAS사와의 협업을 소비자 불만사항을 분석하고 공통이슈를 파악하여 문제 해결의 우선순위를 결정



1 데이터의 개념

2 데이터의 활용

3 데이터와 혁신

4 시사점

시사점

○ 금융기관의 대응*

- › 데이터 기반 의사결정을 내릴 수 있는 환경 구축
- › 시스템의 단순화 / 자동화 / SaaS 도입으로 운영비용 감축
- › Cloud 및 API를 활용하여 언제 어디서든 연결할 수 있는 환경 구축
- › 고객 니즈 파악하기 위한 데이터 수집 및 분석 역량 강화
- › 사이버보안에 대한 관심
- › 관련 인력 및 기술 확보

* Six priorities for 2020, PwC(2016) 참조

시사점

○ 유의사항

- › 한국의 개인정보 규제는 엄격한 편
 - 까다로운 비식별 절차, 다른 데이터와의 결합을 통해 재식별 가능성, 법적 면책 여부 등
- › 개인정보의 무분별한 활용으로 개인권리를 침해하는 사례 발생
 - 고객이 원치 않는 DM 급증, 보이스피싱 등
- › 인과관계 고려 없이 상관관계만 파악하여 과도하게 의미를 부여할 경우 잘못된 판단을 내릴 뿐 아니라 사회적 부작용을 야기할 수도 있음
 - 미국 경찰은 중고차를 운전하는 20~27세 흑인 및 히스패닉 남성이 약물 가능성이 높을 것으로 예측하고 검문을 실시한 바 있음 (안춘모, 2017)

시사점

- 사회문제 해결 위한 빅데이터의 활용 가능성
 - › 의료, 교통 등 타 분야에서 빅데이터가 사회적 현안 해결 실마리 제공
 - › 금융 분야 현안 예시
 - 금융상품 판매 시 발생하는 금융기관-고객 간 이해상충 문제
 - 퇴직연금 자산에 대한 가입자의 무관심 및 방치
 - 개인투자자의 비이성적 거래행태에 따른 수익률 저하
 - 공모펀드의 저조한 수익률 등
 - › 문제의 본질을 파악하고 현상을 개선하는데 중요한 역할 수행



Thank You!