DB Issue Report

데이터 기반 산업 관련 해외 동향

– DDS, DDE, DDI를 중심으로 –

주요 내용

I. 개요
II. 해외 동향
III. 시사점

KODB 한국데이터베이스진흥원
KOREA DATABASE AGENCY

I 개요

o 데이터 기반 산업과 '데이터 주도(Data-Driven)'

- DB산업은 데이터를 축적하고 정보화하는 것에서 벗어나 '데이터를 분석하여 의미를 찾아내고 이를 실행에 연결시키는 산업'이 되고 있음[1)]
- 데이터 기반 산업은 DB산업에서 출발하지만, 데이터에 집중하면서 발생하는 새로운 산업적 활동(빅데이터, IoT, 클라우드 등)에 초점
- 미국 M.I.T. Media Lab의 '데이터 주도 사회(Data-Driven Society)', OECD와 SIIA(Software & Information Industry Association)의 '데이터 주도 혁신(Data-Driven Innovation)', EC의 '데이터 주도 경제(Data-Driven Economy)' 참조

o 'Data Driven'의 의미

- 어떠한 행동(활동)이 직관이나 개인 경험이 아닌 데이터에 의해 일어나게 되는 과정

> 'Data-driven' means that progress in an activity is compelled by data, rather than by intuition or personal experience[2)].

- 그러나 여러 분야에서 약간씩 다른 의미로 사용되기도 함

• Data-Driven Mind-Set	• Data-Driven Environment
• Data-driven programming	• Data-driven Conversations
• Data-driven journalism	• Data-driven marketing
• Data-driven testing	• Data-Driven Advertising
• Data driven learning	• Data Driven World (ENCORE)
• Data-driven science	• Data Driven Future

1) 한국데이터베이스진흥원, 2013 데이터베이스백서, 2013.7.
2) Wikipedia, 'Data-Driven' 검색 결과. http://en.wikipedia.org/wiki/Data_driven (접속일 : 2014.9.25)

Ⅱ 해외 동향

□ M.I.T. Media Lab의 데이터 주도 사회(Data-Driven Society)[3)]

o 배경

- M.I.T. Media Lab Alex Pentland 교수를 중심으로 학계, 산업계, 언론계 관련 인사들이 새롭게 대두되고 있는 '데이터 주도 사회'의 개념 정립을 위한 논의 중
- 데이터 분석은 재무 시스템, 정부 기능, 효과적인 보건의료 등 여러 분야에 활용할 만하며, 특히 빅데이터 분석을 위한 프레임워크를 정립해야 함(프로메테우스의 불)

o 디지털 빵 부스러기의 잠재력

- 우리는 살면서 디지털 빵 부스러기(페이스북 현재 상태, 트윗으로 퍼 나르는 정보들)를 남기게 되며, 이로 인해 자신의 행동을 기록으로 남기게 됨
- 이러한 행동들은 빅데이터의 수학적 패턴이 되어 사회 기능에 통찰력(insight)을 제공(빅데이터로 문제 진단 및 해결)

o 데이터 뉴 딜(A New Deal on Data)

- 데이터 주도 사회를 이루기 위해서는 데이터 뉴 딜(공공재로서 데이터를 이용 가능하게 하면서 동시에 국민을 보호) 필요
- 데이터 뉴 딜의 핵심은 개인 데이터(personal data)를 하나의 자산으로 다루어 개인이 데이터 오너십을 발휘할 수 있도록 하는 것(옵트인 방식의 개인정보 보호)

3) TheNewYorkTimes, The Promise and Peril of the 'Data-Driven Society', 2013.2.25. 및 Alex 'Sandy' Pertland, The Data-Driven Society, Scientific American 309, 2013, p.78~p.83. 참조.

o 리빙 랩(living labs)

- 사회 시스템이 너무 크고 복잡하기 때문에 합리적 가설이 매우 많아지고 관련 데이터를 모두 분석하는 것은 불가능
- 이에 따라 데이터 주도 사회를 건설하기 위해 아이디어를 테스트해 볼 수 있는 "리빙 랩"을 만들어야 함

<참고> 리빙 랩 사례 : 열린 데이터 도시(open-data city), 이탈리아 트렌토(Trento)

- 트렌토 시와 통신사, 대학 연구소, 데이터 주도 디자인 연구소 등이 협동으로 열린 데이터 시 프로젝트를 시작
- 트렌토 내에 보다 나은 아이디어가 흘러다닐 수 있도록 진흥시키려는 목적
- 열린PDS(Personal Data Store)를 설치하여 데이터 뉴 딜을 실천(의료 정보, 자녀 육아 정보 등의 개인 정보 공유시 보안 강화)
- 개인이 개인정보 공유에 동의하면 열린PDS를 통해 안전하게 익명화 처리되며, 다른 사람들과 자동으로 정보 공유

□ EC의 데이터 주도 경제(Data-Driven Economy)[4)]

o 배경

- 전세계적으로 빅데이터 기술과 서비스가 성장하는 추세로 빅데이터는 다양한 분야에서 엄청난 잠재력을 보유하고 있음
- 그러나 미국에 비해 유럽의 디지털 경제는 데이터 혁명을 수용하는데 뒤쳐져 있으며, 성공한 데이터 회사도 부족한 편임

※ 데이터 전문가 부족, 대용량 데이터에 대한 엑세스 권한 미약, 복잡한 법제도 환경과 인프라 구조가 중소기업의 진입 장벽으로 작용하여 혁신을 방해

4) European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Towards a thriving data-driven economy, 2014. 2. 7.

- 글로벌 지배자가 아직 등장하지 않은 초기 단계이므로 새로운 기회를 포착, 데이터 경제의 세계적 경쟁력을 확보하기 위한 대책 필요

o 미래 지식경제 및 지식사회의 중심, '데이터'

- 데이터 분석은 보다 나은 결과, 처리, 의사 결정을 가능케하는 수단이며, 새로운 아이디어나 해법을 제시하고, 미래를 보다 정확히 예측하게 함

- 데이터 중심 혁신(Data-driven innovation)은 새로운 일자리를 창출하나 데이터 분석과 자동 학습, 시각화 뿐만 아니라 데이터 오너십, 라이선스 제한, 데이터 보호 등의 법률 분야에 이르는 범학문적 전문가 필요

o 미래 데이터 주도 경제 육성 방안

계획안			내 용
커뮤니티 조직화	데이터 분야 유럽 공공-민간 파트너십		● cPPP(contractual Public-Private Partnership)을 통한 전략적 협력관계 조성 ● 지식과 기술의 이전 활성화를 위한 파트너십 및 체계
	디지털 기업 정신과 오픈 데이터 인큐베이터		● 기업 대상 디지털 기술의 무한 잠재력 인식 제고 ● 오픈 데이터 인큐베이팅으로 중소기업이 데이터를 기반으로 공급망을 형성, 공개적이고 공정하게 데이터 자원을 활용할 수 있는 환경 조성
	기초 기술 육성		● 고급 데이터 전문가 양성을 위한 유럽 경쟁력 네트워크 조성
	데이터 시장 모니터링 도구		● 유럽 데이터 시장의 크기와 동향, 주체 간 상호작용을 파악하기 위한 모니터링 도구 개발
	분야별 연구 투자 우선순위 선정		● 건강, 에너지, 환경, 사회과학과 통계 등 연구 커뮤니티 조성 및 공공-민간 펀딩 마련
프레임워크 상태 개선	데이터 가용성 및 상호 운용성	데이터 개방 정책 활성화	● 데이터 개방 정책 및 법률 프레임워크를 위한 표준 라이선스, 데이터셋, 재이용 과금 등에 관한 가이드라인 마련 ● EU 오픈 데이터 포탈과 원스탑 샵을 위한 범유럽 오픈 데이터 디지털 서비스 인프라 마련
		데이터 처리도구 및	● 중소기업 및 웹 기업 대상 비즈니스 인텔리전스, 의사결정 지원 시스템 등 연구 혁신 지원(데이터 분석, 시각화, 인공 지능, 의사결정 지원

		방법론 개발	SW 및 알고리즘 등) ● 클라우드 기반 데이터 인프라(PaaS, SaaS) 지원 ● 중소기업 경쟁력 지원 강화를 위한 네트워킹 강화(개발, 데이터 접근, 기술 및 서비스, 비즈니스 절차 및 활동 개선 등)
		새로운 개방 표준 개발 지원	● 주요 경제분야 전 유럽 대상 표준 개발 ● 데이터 개방 및 교환을 위한 분위기 조성, 빅데이터 관련 현재 표준의 맵핑 지원
	데이터 중심 경제를 위한 인프라 개선	클라우드 컴퓨팅	● 투명한 평가 기준, 전 유럽의 자발적 인증, 안전하고 공정한 약관, 클라우드 이용자의 환경 개선을 위한 클라우드 컴퓨팅 전략 ● 데이터 중심 경제 활성화를 위한 유럽 클라우드 파트너십(European Cloud Partnership) 마련
		e-인프라스트럭처 고성능 컴퓨팅	● HPC(High Performance Computing)을 과학·산업·사회 문제 해결에 적용할 수 있도록 추진 조직 설립(cPPP 적용)
		네트워크/브로드밴드/5G	● 5G 분야 cPPP를 통한 미래 모바일 인터넷 기술 개발, 개인 투자 유치를 위한 규칙과 펀드 모집
		사물 인터넷	● 사물 인터넷을 통한 데이터 수집시 발생하는 데이터 가용성, 품질, 상호운용성 문제 해결을 위한 대규모 자금 확보 프로젝트
		공공 데이터 인프라	● 상호 협력과 효율성 증대를 위한 데이터 처리 기관 간 내부 네트워크 연결(지역 데이터 센터 연계, 중소기업-학계-연구계-공공기관 연계)
	법제화	개인정보 보호, 소비자 보호	● 디지털 환경에서 개인의 신뢰와 법적 안전성을 확보할 수 있는 법규 환경 조성 ● 데이터 익명성, 가상성, 데이터 최소화, 개인 데이터 위험도 분석 등 주요 이슈에 관한 가이드라인 마련 ● 이용자 통제 클라우드 기반 기술, 개인 데이터 사용, 데이터 공유 정책에 필요한 이용자 지원 툴 개발 지원을 위한 컨설팅 수행
		데이터 마이닝	● 데이터 마이닝 기반 데이터 중심 혁신 방안 조사(저작권 관련 사항 포함)
		보 안	● 빅데이터 관련 보안 위협 요소 조사 및 위험 관리·완화 조치, 가이드라인, 데이터 스토리지 보안 성공 사례 등 조사 ● 데이터 침해·데이터베이스 침해 위험 감소를 위한 연구 혁신 지원
		데이터 오너십과 이전	● 데이터 이전시 발생하는 장애 조사 및 정책적 조치 마련 ● 데이터 오너십 및 데이터 제공시 책임 소재, IoT를 통한 데이터 수집시 문제들에 대한 가이드라인 필요성에 대한 평가 컨설팅 실시

□ OECD의 데이터 주도 혁신(Data-Driven Innovation)[5)]

o 배경

- 사회·경제 활동에 대한 마이그레이션 증가와 데이터 수집·전송·저장·분석 비용의 하락, 데이터량의 폭발적 증가
- 데이터는 사회·경제적 가치로 인해 더욱 중요해지고 있으며, 이는 데이터 주도 사회경제모델(data-driven socioeconomic model)로 전환 중

o 5대 분야의 데이터 혁신과 생산성 향상

- OECD에서 사전 연구를 통해 데이터 활용시 혁신과 생산성 향상 가능성이 높은 5대 분야를 선정

분야	내 용
온라인 광고 (online advertisement)	● 클릭 순서 데이터(Click Stream Data)로 소비자의 인터넷 사용 습관 분석 ● 2012년 상반기 미국 상위 5백개 온라인 광고회사의 매출액은 84억 달러(8조 7천억원)로 2011년 상반기 대비 15% 성장
보건·의료 (health care)	● EHRs(Electronic Health Records)의 통합으로 기록 관리 비용 절감, 의료 사고 감소, 진단 및 처방 개선, 환자의 의료진·의료 설비 선택 용이, 의료와 보험 연계 강화 등 효과 ● 미국은 2020년까지 약 3천억 달러(312조원) 절감 예상
설비 (utilities)	● 스마트 그리드 기술로 전력 소비 패턴을 감지하여 전력을 절감(스마트 미터는 실시간 전력 소비량과 요금을 표시해 주며, 주택 내의 다양한 전자 기기를 켜고 끌 수 있도록 제어) ● 2020년까지 이산화탄소 20억 톤(790억 유로, 105조원) 감축 효과
물류·교통 (logistics and transport)	● 모바일 기기의 증가로 위치 추적이 수월해 지면서 교통 상황 모니터링과 교통 시간 단축이 가능해 짐 ● 2020년까지 연료 절감, 이산화탄소 배출 감소 등으로 약 5천억 달러(520조 7천억) 절감 예상
공공행정 (public administration)	● 날씨예보, 교통관리, 범죄통계, 정부 투명성 개선, 교육 및 문화 지식 등 공공 데이터 활용으로 행정 비용 절감 가능 ● 유럽 23개국의 경우 15~20% 가량 절감(15백억 유로, 2백조원) 가능할 것으로 예측 ● 2010년 유럽 공공정보 활용 시장 규모는 약 320억 유로(42조 6천억원)

5) OECD, Exploring Data-Driven Innovation as a New Source of Growth - Mapping the Policy Issues Raised by "Big Data", 2013. 6. 18.

o 정책 함의

- 빅데이터를 이용한 혁신을 위해서는 데이터 수집, 전송, 저장, 가공, 이용에 관한 정책과 관행을 다음과 같이 개선해야 함

분야	내 용
사생활 보호 Privacy Protection	● 인터넷 경제 속에서 신뢰와 혁신이 가능한 사생활 보호 ● 효과적인 사생활 보호와 개인 데이터 이용에 따른 경제사회적 편익 간의 타협점 필요
데이터 접근성 Open Access Data	● 분야를 초월한 데이터의 연계와 활용이 있어야 혁신과 사회경제적 발전이 가능(익명 처리된 모바일 트래픽 데이터와 자동차의 네비게이션 시스템, 공공 도로 관리 시스템) ● 아직까지 데이터 공유가 많이 이루어지지 못하였으며, 데이터 공유의 경제적 인센티브도 충분히 보장받지 못하는 상황
인력 (재직자 능력개발) Employment	● 데이터 관리 및 분석 능력 보유자(데이터 과학자)의 수요와 공급 불일치 발생 ● 이로 인해 빅데이터 분석 시도가 저해되고 있으며 새로운 일자리 창출의 기회 상실 ● STEM(Science, Technology, Engineering, Mathematics) 분야 교육, 훈련, 기술 개발을 위한 범학문적 접근법 필요
인프라 Infrastructure	● 수 많은 기기 간의 인터넷 연결 ● IPv6 인터넷 주소 체계 정비, 기업 대상 모바일 시장의 접근 채널 확대, 한정된 자원으로 최대의 이익을 낼 수 있도록 하는 주파수 할당 규제 등
평가 Measurement	● 아직까지 경제통계상 데이터의 가치는 저평가되어 있음 ● 정부와 기업이 빅데이터 분석을 통해 잠재된 편익과 위험 요인을 찾아낼 수 있도록 적절한 정책을 마련해야 함 ● 데이터 활용 촉진의 위험 요인을 줄이고 편익을 확대할 수 있는 평가 필요

□ SIIA의 데이터 주도 혁신(Data-Driven Innovation)[6]

o 배경

- 데이터 주도 혁신(DDI)은 우리가 일하고, 소통하고, 학습하며 우리의 삶을 살아가는 방식을 변화시킬 수 있는 엄청난 경제적·사회적 가치 제공
- 본 보고서는 정책 결정자를 위한 지침서로, 데이터에 대한 이해 뿐만 아니라 데이터를 통해 경제적·사회적 가치를 활성화할 수 있는 로드맵임

o 정책 입안자 및 정부를 위한 제언

- 정책 제언의 기본 원칙은 ① 데이터 수집·분석을 제약하는 광범위한 정책을 입안하지 않도록 하는 것 ② 정책을 현재의 '스냅샷(snapshot)'을 기반으로 만들지 말고, 장기적 발전을 감안하여 의무보다는 유연성하고 개방적인 규정을 만들 것
- 10가지의 권고 사항

1. (DDI의) 잠재력을 최대한 발휘하기 위해 위험 요소와 사회적 이익을 모두 고려한 사생활 보호 발전 전망이 담긴 데이터 주도 혁신 정책 프레임워크를 마련해야 한다.

- 정책 입안자들은 사생활 보호 법규와 잠재된 사생활 침해 위협 간 균형 유지
- 사회통념상 용인되는 사생활 보호 기준이 기술과 함께 진화함을 인식할 것

2. '데이터 최소화 원칙'은 DDI의 입장에서 재해석해야 한다.

- '데이터 최소화 원칙'이 모든 법제도에 형식적으로 엄격히 적용되어서는 안되고 보다 나은 데이터 관리를 위해 위험 요소와의 균형이 필요
- 사생활 보호 기술과 데이터 원칙 준수를 통해 사생활 보호와 침해 위협의 균형을 보장하는 효과적인 데이터 최소화 원칙 필요

6) SIIA, Data-Driven Innovation A Guide for Policymakers: Understanding and Enabling the Economic and Social Value of Data, 2013.

3. DDI와 사생활 보호 사이에서 형평성을 고려한 개인 익명(비식별)화를 장려해야 한다.

- 개인 식별 정보가 없더라도 DDI의 주요 결과물을 도출할 수 있음
- 개인정보가 수집되더라도, 주요 공공·사회적 목표 달성을 위한 가치나 효용에 영향을 미치지 않는 방법으로 즉시 익명화 가능

4. 개인정보 수집 및 (수집) 동의에 똑같은 형식을 광범위하게 적용해서는 안된다.

- 정책 입안자들은 고지에 입각한 (정보 수집) 동의의 실상과 개인정보의 수집에 대한 명시적 동의가 필요한 특정 상황에 대해 지속적으로 고려해야 함

5. 정책 입안자들은 기술 중립성을 유지하고 기술의 의무 사용을 지양해야 한다.

- 기술 중립성은 인터넷 기반 IT산업의 정책 원리로 널리 인정되고 있으며, 기술의 의무 사용에 대한 저항, 기술 중립 유지를 위해 지속된 이러한 원칙은 DDI 생태계에 특히 중요
- 보다 나은 결과와 사생활 보호가 가능한 보다 빠르고, 보다 나은, 비용이 덜 드는 혁신적 방법을 모색할 수 있는 정책 필요

6. 개방형 표준은 DDI의 중요한 성공 요인으로 정부 아닌 산업계가 주도하는 표준 개발 기관을 통해 계속 발전해야 한다.

- 여러 데이터를 통합하는 능력은 DDI의 핵심 요소이며, 다양한 분석 환경과 애플리케이션을 모두 포괄하는 광범위한 데이터셋의 결합에 개방형 표준이 필요
- 정부는 데이터에 개방형 표준을 적용하도록 장려할 수 있으나 특정 기술 표준에 (공인) 자격을 부여하거나 특정 분야에 새로운 표준을 직접 만들려 해서는 안됨

7. 데이터 수집·통제 주체가 데이터 관리·분석 서비스 제공자와 함께 작업할 경우, 사법권역이 서로 다른 경우에도 사생활 보호 및 보안 규정이 상이하지 않도록 조율하는 정책이 필요하다.

- 데이터 주도 혁신으로 고객에게 가치를 제공한다는 조건으로 다양한 이해

당사자들이 연관되어 있음. 개인정보 보호 및 보안 규정과 계약은, 데이터가 관할 권역을 넘어가는 경우에도, 데이터 주체를 보호하기 위해 공조되어야 함

- 정부는 데이터 관리·분석 서비스 기업들이 경계를 넘어 글로벌 기업이 될 수 있도록 데이터 정책 프레임워크에 상호운용성을 갖춰야 함

8. 학생들의 사생활 보호 수요에 대응하면서 동시에 DDI로 인해 교수 학습 수준이 대폭 향상될 수 있는 정책이 필요하다.

- 학생, 가족과 학교는 학생 개개인별 맞춤형 학습을 제공하기 위해 다양한 기술을 사용하는 추세
- 기술 기반 교육 제품·서비스 제공자는 점점 우수 학습 환경을 제공하기 위해 데이터를 분석하고 있으며 보다 효율적인 학습 방법을 통해 비용을 절감할 수 있어 학생과 교육기관은 혜택을 받음
- 이러한 서비스 제공자들의 역량을 약화시키지 않으면서 어린이들의 사생활을 충분히 보호할 수 있는 균형있는 정책이 필요

9. 정부는 보다 효율적이고 효과적인 정부, 예산 낭비를 줄일 수 있는 정부가 될 수 되도록 DDI를 촉진시키는 정책을 채택해야 한다.

- 오늘날 정부는 각계로부터 생산성을 향상시키고 보다 나은 시민 서비스를 제공하면서 운영 비용을 절약하도록 압박을 받고 있음
- 이에 정부는 객관적 데이터 분석을 토대로 정책을 채택해야 하며 또 데이터 과학 연구 개발을 장려하며 데이터 과학자와 전문가를 활용할 수 있어야 함

10. 정부는 주요 공공데이터에 대한 접근성을 극대화하는 데이터 개방 정책과 공공-민간 파트너십을 정책에 지속적으로 반영해야 한다.

- 정부는 귀중한 데이터를 대량 보유하고 있으며, 국민들은 정부 데이터에 대한 접근을 원하고 있음
- DDI의 잠재력을 극대화하는 효과적인 방법은 정부가 데이터 공유가 가능한 기업 아키텍처를 채택하고, 주요 공공 데이터에 대한 접근권을 제공하기 위해 공공-민간 파트너십을 활용한 데이터 개방 정책을 수용하는 것

III 시사점

o 국제기구와 각국 정부는 '데이터'에 주목하기 시작

- OECD와 미국, 유럽을 중심으로 데이터가 주도하는 사회적 변화, 특히 빅데이터로 촉발된 데이터의 가치, 그리고 그 효과를 증폭시키기 위한 정부 정책에 대한 재조명 작업이 활발히 진행 중

- 그러나 정책 내용은 '개인정보보호', '인력 양성', '인프라 확충', '기초 기술 개발', '공공 데이터 개방' 등으로 국내 빅데이터 활성화 정책과 유사하며 빅데이터 이전 정책과도 유사

o 빅데이터 활성화 정책의 방향

① 데이터 주도형 의사 결정을 국가 전략의 핵심으로 설정

- 소규모의 우수 데이터 과학자, 엔지니어로 구성된 팀을 만들고 의사결정권자와 양방향의 지속적인 커뮤니케이션 채널 설정(빅데이터 분석의 목적, 즉 해결해야할 문제 부터 명확히)

② 데이터 거버넌스 체계 수립

- 데이터에 집중한 관리 정책, 지침, 표준, 전략 및 방향, 범정부 차원에서 데이터를 관리할 수 있는 조직과 서비스를 우선 마련해야 함

③ 성공 사례의 수집과 전파

- 성공한 빅데이터 프로젝트의 분석 노하우를 전파, 다른 문제에 대한 적용 가능성 확인, 적용 과정 모니터링 및 개선 필요

④ 전문가 육성과 협업 체계 마련

- 데이터 과학자 육성을 위한 실습 중심 훈련 프로그램을 마련해야 하며, 다양한 데이터를 활용하기 위한 부처 간 협업 체계 필요

- 본 자료를 인용하실 경우 출처를 반드시 명시해 주십시오.
- 내용에 대해 문의사항이 있으신 경우 아래로 연락해 주십시오.

연락처 : 02-3708-5361, taehoon@kodb.or.kr