

빅데이터의 동향 및 시사점

■ 김 한 나*

1. 개요

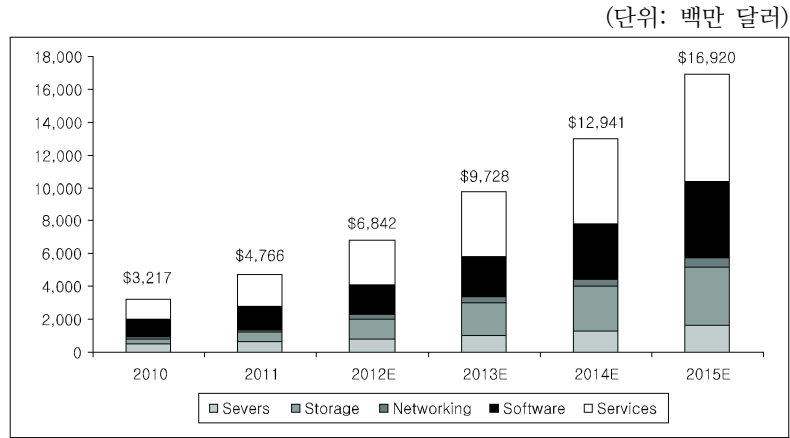
최근 ICT 분야에서 ‘빅데이터’ 이슈가 급부상하고 있다. 디지털 정보량이 기하급수적으로 증가함에 따라 수많은 데이터를 어떻게 활용하는지의 여부, 즉 방대한 데이터를 통한 새로운 가치창출이 기업뿐 아니라 국가의 경쟁력 강화와 직결되는 시대로 접어들고 있는 것이다.

미국 MIT 슬론스쿨 조사에 따르면, 모바일, 온라인 상거래, 소셜네트워크 서비스 등에서 하루에 250경(2.5×10^{18}) 바이트 분량의 비정형 데이터가 발생하며, 매달 300억여 개의 페이스북 메시지와 10억여 개의 트윗이 생성된다.¹⁾ 시장조사 전문기관인 IDC는 ‘전 세계 빅데이터 기술 및 서비스 전망보고서’에서 2011년 한해 동안 새롭게 생성되거나 복제된 정보의 양이 1.9제타바이트(1조 9천억 기가바이트)를 넘어섰고, 향후 5년 내 거의 9배까지 증가할 것이라고 예측했다. 또한 전 세계 빅데이터 시장은 2010년 32억 달러에서 오는 2015년에는 169억 달러 규모에 달할 것이며, 연평균 40% 성장 및 전체 정보통신기술(ICT) 시장 성장률의 약 7배에 달할 것으로 전망했다.²⁾

* 정보통신정책연구원 통신전파연구실 연구원, (02)570-4123, khn37@kisdi.re.kr

1) 《매일경제》(2011. 11. 4).

〈표 1〉 빅데이터 세계시장 규모



구분	2010	2011	2012	2013	2014	2015	CAGR(%)
서버	495	665	803	1,032	1,270	1,657	27.3
스토리지	318	560	1,224	1,968	2,719	3,479	61.4
네트워킹	106	146	242	368	485	620	42.4
S W	1,062	1,415	1,851	2,476	3,376	4,625	34.2
서비스	1,236	1,979	2,721	3,883	5,099	6,538	39.5
합계	3,217	4,766	6,842	9,728	12,941	16,920	39.4

자료: IDC(2011)

이처럼 빅데이터가 차세대 ICT를 이끌어 갈 핵심동력으로 주목받는 가장 주된 요인은 기존과 차별화된 대용량 데이터의 새로운 분석과 추론(전망)을 통해 새로운 서비스를 개발할 수 있는 가능성이 무궁무진하기 때문이다. 대용량 데이터를 기반으로 구글의 ‘자동번역’ 및 애플의 ‘시리’와 같은 자연언어 처리, 기계 학습, 인공지능 기술로 맥락 이해와 추론 서비스가 증가되고 있다. 그리고 이러한 추론의 영역은 의료, 금융 등의 공공 분야를 혁신하는 新산업 분야를 개발하고, 맞춤형 개인화 서비스를 가능하게 한다.

기업들은 이러한 데이터를 활용, 수집, 분석하여 소비자의 마음을 읽는 ‘소셜 애널리틱스(Social Analytics)³⁾’를 마케팅 수단으로 활용하고 있으며, 각 국가들은 행정, 의

2) 《아이뉴스》(2012. 3. 29)

료, 복지, 교육 등의 공공 분야에서 빅데이터를 정책 수립 및 국정운영의 원천으로 활용하고 있다.

이와 관련하여 본고에서는 빅데이터의 국내외 활용 사례와 빅데이터와 관련된 시사점을 살펴보고자 한다.

2. 빅데이터 활용 사례

(1) 공공·행정 분야

세계 각국의 정부들은 빅데이터가 미래사회를 위한 새로운 경제적 가치의 원천이 될 것으로 기대하고, 각 분야에 걸쳐 데이터 분석에 의한 예측과 서비스 개발을 위해 노력하고 있다. 데이터 분석을 통해 사회현상에 대한 미래 통찰력을 얻을 수 있으며,

〈표 2〉 국가별 빅데이터 관리 및 활용 방법

국 가	추진기관	세부내용
미 국	대통령 직속 과학기술정책실	- '빅데이터 연구개발 이니셔티브' 발표(2012. 3) - 부처별, 지방정부별 빅데이터 활용한 서비스 발굴 및 운영 - 공공정보 데이터 공개사이트 'data.gov' 운영
영 국	기업혁신기술부	- '데이터 전략위원회(Data Strategy Board)' 설립(2012. 3) - 데이터 공유플랫폼 'data.gov.uk' 운영
싱가포르	경제개발청	- 국가안보조정국(NSCS) 내 RASH 시스템 마련(2004. 7) - 민간협력으로 데이터분석연구소 설립 - 공공정보 데이터 공개사이트 'data.gov.sg' 운영
대한민국	대통령 직속 국가정보화전략위	- '빅데이터를 활용한 스마트정부 구현방안' 마련(2011. 10) - 빅데이터 마스터플랜 추진 및 빅데이터 국가전략포럼 실시 - 공유자원포털 'data.gov.kr' 운영

자료: 국가정보화백서(2012) 재구성, pp.53~56

- 3) 소셜 애널리틱스: 소셜 네트워크서비스(SNS)상의 대량 메시지를 신속하게 분석하는 기술로, 사람의 감정을 나타내는 단어가 나오면 앞뒤 문맥에 따라 긍정적 혹은 부정적 반응인지 판단하고, 두 단어 이상을 조합해 걸러내는 텍스트 분석 등을 활용함

위험징후 및 이상신호를 포착할 수 있기 때문이다. 국가별 빅데이터 관리 및 활용 상황은 앞의 <표 2>와 같다.

1) 미국

미국은 대통령 자문위원회인 과학기술정책자문위원회의 ‘Designing a Digital Future’ 보고서(2010. 3)를 통해 대규모 데이터의 가치에 주목하고, 모든 연방정부에 빅데이터 전략을 수립할 필요성이 있음을 제시하였다. 이후 2012년 3월, 대통령 직속기관인 과학기술정책실(Office of Science and Technology Policy: OSTP)은 빅데이터 연구개발 이니셔티브를 발표하고, 2억 달러 이상을 투입해 빅데이터 기술개발을 위한 계획을 마련하였다.

<표 3> 미국의 빅데이터 이니셔티브 추진 현황

중앙부처 및 지방정부	세부내용
국가테러방지센터 (NCC)	- SNS, 신문, 잡지 등 매일 1만 건 이상의 테러 관련 정보 수신 및 데이터 분석 - 테러동향 파악 및 테러징후의 사전예측 체계 마련
유전자정보은행 (CODIS)	- 1,100만 건 이상의 범죄자 유전자 정보 분석 - 범죄용의자 추적에 활용하여 18만 건의 검거에 활용
국립과학재단 (NSF)	- 데이터 관리, 분석, 시각화, 지식 추출 등을 위한 핵심연구, 기술개발 지원 및 인력양성 - 빅데이터를 이용한 지식 추출 방법론을 위한 장기적인 전략 수립 - 연간 투자규모: UC.Berkeley(1,000만 불), 인력양성기관(200만 불), 연구그룹(140만 불) 등
국방부 (DoD)	- 데이터를 이용한 국방 의사결정 체계 구축 - 별도의 빅데이터 관련 예산 매년 2억 5천만 불 투입
국립보건원 (NIH)	- 75개 기업 및 기관 파트너십을 통해 ‘1,000 유전체 프로젝트’ 추진 - 유전적 데이터를 통한 유전적 다양성 분석, 난치병 치료, 신약 개발 - 필박스(Pillbox)의 국민의 약 검색통계를 활용하여 AIDS 등 관리대상 주요 질병의 분포, 연도별 증가추세 등을 예측
방위고등연구계획국 (DARPA)	- 대용량 데이터 분석을 위한 고성능 컴퓨팅/SW 기술개발 - 다양한 미션에 쉽게 적용 가능한 HCI 기술개발 - XDATA 프로젝트(연간 2,500만 불)

중앙부처 및 지방정부	세부내용
에너지부 (DoE)	- Advanced Computing을 통한 과학적 발견 - 로렌스버클리 국립연구소(2,500만 불)
지질조사원 (USGS)	- 지구 시스템 과학을 위한 빅데이터 기술
미시간 주	- 정부기관 통합 데이터웨어하우스(Integrated Data Warehouse, IDW) 구축, 공공의료보험 부정행위 발생 감지, 개인 건강관리 개선 등 서비스 품질 개선
오하이오 주 오클라호마 주	- 국세청(IRS) 데이터와 고용데이터로 마이닝(mining) 분석을 통해 새로운 세원과 미납세금 확인

자료: 국가정보화백서(2012) p.53, 안창원, 황승구(2012) p.11, 이강용 외 4인(2012) p.42, 정용찬(2012) p.10 재구성

이처럼 미국은 대통령 중심의 빅데이터 이니셔티브를 마련하여 추진하는 동시에, 부처별 차원에서 다양한 정책을 발굴 및 추진하고 있다. 미국 오바마 정부는 ‘투명하고 열린 정부’라는 국정운영을 발표하고, 교통, 통신, 의약품, 안전, 범죄 등의 공공정보 데이터를 ‘data.gov’를 통해 제공하고 있다. 현재 총 39만 2,498개의 데이터 세트를 제공하고 있으며, 1,345개의 애플리케이션이 개발되어 공공정보를 소비자의 의도와 취향에 맞게 사용 가능하도록 서비스를 제공하고 있다.⁴⁾

2) 영국

영국은 ‘영국 역사상 가장 투명한 정부’를 목표로 오픈 데이터 전략을 추진하고 있다. 기업혁신기술부(BIS)는 공공정보 공개 및 데이터 가치창출을 위한 자문기구인 ‘데이터 전략위원회’를 설립하고, 내각사무처(Cabinet Office)와 함께 데이터 접근성 강화 및 데이터 공개지침 등의 정책을 발표하였다. 영국도 데이터 공개를 위해 ‘data.gov.uk’의 데이터 공유 플랫폼을 재정비하여 운영하고 있으며, 현재 8,702개의 데이터 세트를 제공하고 있다.⁵⁾ 대표적인 빅데이터 활용 사례인 영국 국민의료서비스(NHS)는 전국 약국, 병원의 처방 데이터를 기반으로 특정 지역, 특정 질병 가능성을 분석하고 있다.

4) 미국 공공데이터 공개 사이트 data.gov 참조(2012. 10. 3. 기준)

5) 영국 공공데이터 공개 사이트 data.gov.uk 참조(2012. 10. 3. 기준)

3) 싱가포르

싱가포르는 2004년 7월, 총리실(Prime Minister's Office) 산하 국가안보조정서사무국(NSCS: National Security Coordination Secretariat) 내에 국가위기관리시스템(RASH: Risk Assessment and Horizon Scanning)을 마련하여, 다양한 데이터를 수집·분석하여 국가적 위험 및 기회에 대한 사전예측 및 대응방안을 모색하고 있다. 그리고 2012년에는 RASH Program Office를 설립하여 업무의 체계적 수행을 위한 환경 및 이슈 분석, 정책수립 능력 강화, 시스템 기술 연구에 특성화된 3개의 센터를 운영하고 있다. 경제개발청(Economic Development Board)은 민간협력을 통해 데이터 분석연구소를 설립하여 데이터 분석을 기반으로 한 정책수립체계를 마련하였고, 2011년 6월에는 공공데이터 오픈사이트인 'data.gov.sg'를 개설하여 현재 50여 개 행정기관의 5,000여 개의 데이터 세트를 제공하고 있다.⁶⁾

4) EU

EU는 '데이터 개방전략(ODS: Open Data Strategy)' 차원에서 2011년 12월에 유럽연합기구와 27개 회원국에서 생산하는 모든 공공정보와 데이터를 의무적으로 공개하고, 사용자들이 이러한 데이터들을 자유롭게 활용할 수 있도록 EU집행위원회 지침을 개정했다. 이 지침에 따르면 2013년 'EU 데이터 포털'이 출범할 때까지 EU 공공기관 및 27개 회원국 부처, 공립대학, 도서관, 박물관 등의 공공데이터 관리기관들은 문서자료를 비롯한 지도, 지질학적 자료, 이미지, 동영상, 웹 애플리케이션 등의 공공데이터 자료를 온라인으로 게재해야 한다. 또한 일반 시민과 기업은 이와 같은 공공데이터를 상업적으로 재가공한 뒤 부가가치를 높여 판매할 수 있다. EU는 공공데이터를 기반으로 정보 가공산업이 확대되면 경제가 활성화되고, 일자리가 늘어나 연간 400억 유로(약 61조 원)에 이르는 생산유발 효과를 가져올 것으로 예측하고 있다.⁷⁾

6) 싱가포르 공공데이터 공개 사이트 data.gov.sg 참조(2012. 10. 3 기준)

7) 《매일경제》(2011. 12. 13)

〈표 4〉 EU 공공정보 공개 의무화 주요 내용

구분	세부 내용
공개 방식	- 공공정보 및 데이터 온라인 게재 의무화 - 개인, 기업정보 재가공 후 판매 허용 - 정보 공개 여부 모니터링할 감시기구 설치
대상 기관	EU 공공기관, 27개 EU회원국 부처, 공립대학, 도서관, 박물관, 기록보관소 등
대상 자료	문서자료, 통계수치, 지도, 지질학적 자료, 이미지, 동영상, 애플리케이션 등
발효 시점	2013년(EU 데이터 포털 출범 예정)
기대 효과	경제활성화와 일자리 창출을 통해 연 400억 유로 생산 유발

자료: 《매일경제》(2011. 12. 13)

〈표 5〉 각국의 공공정보 데이터 공개 현황(2012. 10. 3 기준)

국가명	사이트명	데이터 세트	국가명	사이트명	데이터 세트
미국	data.gov	392,498	대한민국	data.gov.kr	441
캐나다	data.gc.ca	13,069	네덜란드	data.overheid.nl	351
영국	data.gov.uk	8,702	핀란드	suomi.fi	266
싱가포르	data.gov.sg	5,000	사우디	saudi.gov.sa	247
뉴질랜드	data.govt.nz	2,261	그리스	geodata.gov.gr	194
독일	portalu.de	1,630	오스트리아	data.wien.gv.at	187
호주	data.gov.au	1,120	벨기에	data.gov.be	105
덴마크	data.digitaliser.dk	749	프랑스	opendata.paris.fr	75
스페인	datos.gob.es	655	노르웨이	data.norge.no	35

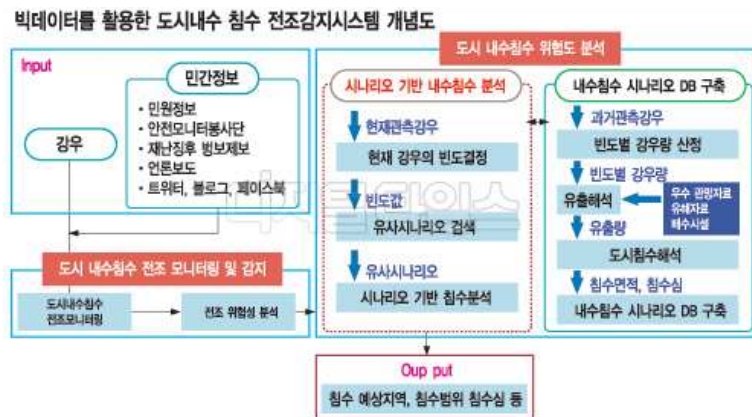
자료: data.gov 외 17개 공공데이터 사이트

5) 대한민국

한국 정부 역시 데이터가 중요한 가치창출의 원천임을 인지하고, 공공데이터 공개 및 빅데이터 마스터플랜 구축 등의 지속적인 노력을 기울이고 있다. 이러한 노력의 일환으로 2011년 10월에는 빅데이터를 통한 정부혁신과 국가경쟁력 제고를 위해 대통령 직속의 국가정보화전략위원회에서 ‘빅데이터를 활용한 스마트 정부 구현방안’을

마련하였다. 또한 행안부, 방통위, 한국정보화진흥원(NIA)은 빅데이터 마스터플랜 T/F 구성 및 빅데이터 분석 및 공공부문 적용 전략의 마련을 위해 2012년 6월에 전문가 자문회의 및 국가전략 포럼을 개최하였으며, 이를 토대로 2012년 8월에 마스터플랜을 수립하였다.⁸⁾ 현재는 이러한 마스터플랜을 기반으로 서울시 내수침수 등 재난안전 분야에서 ‘파일럿 프로젝트(시범사업)’을 추진 중이며, 시범사업을 통해 빅데이터 활용 가능성을 검증할 계획이다.⁹⁾

[그림 1] 빅데이터를 이용한 파일럿 프로그램(시범사업) 개념도



자료: 국가정보화전략위원회(2012), p.9

또한 한국 정부도 미국, 영국 싱가포르와 같이 공공정보 공개사이트인 ‘공유자원포털(data.gov.kr)’을 통해 중앙부처 및 지자체의 공공데이터를 공개하고 있으며, 현재 133개의 서비스와 441개의 데이터 세트를 제공하고 있다.

정부 차원의 빅데이터 활용방안 외에 각 부처 및 공공기관에서도 빅데이터를 적극적으로 활용하고 있다. 방송통신위원회는 ‘빅데이터 서비스 활성화 방안’을 마련하여 7대 과제를 추진 중에 있으며, 국민권익위원회는 2011년부터 각 부처의 온라인 민원

8) 국가정보화전략위원회(2012)
9) 상동, 《디지털타임즈》(2012. 8. 10)

정보를 수집, 관리, 분류, 분석 및 예측하는 민원정보분석시스템을 구축하여 운영하고 있다. 민원정보분석시스템은 월별, 지역별 주요 민원 캘린더를 작성하여 사회적 이슈를 민원지도 형태로 제공하고, 반복 민원에 대한 대책을 마련하기 위하여 주요 이슈에 대한 국민의 의견을 자동분석하고, 경보 및 예보하는 체계를 갖추고 있다.¹⁰⁾

〈표 6〉 방통위의 ‘빅데이터 서비스 활성화 방안’ 7대 과제

과제명	주요 내용
① 신규 서비스 발굴·확산을 위한 시범 서비스 추진	방송통신, 교육, 교통, 의료 등 여러 분야의 혁신적인 시범 서비스를 공모, 발굴
② 빅데이터 기술 및 플랫폼 경쟁력 강화	빅데이터 분석에 필요한 핵심기술 개발 및 오픈소스 기반의 플랫폼 개발
③ 전문인력 양성	대학과 R&D 및 시범사업 연계 추진, 국내외 기업과 제휴하여 교육 실시 및 빅데이터 분석가 자격증 제도 도입
④ 빅데이터 지원센터 설치·운영 및 정보공유 체계 마련	빅데이터 지원센터는 테스트베드, 시범 서비스 수행, 중소기업 컨설팅 제공
⑤ 데이터 산업실태 조사	빅데이터 산업 및 활용실태 파악을 위한 조사 실시
⑥ 개인정보보호 관련 법제도 개선	프라이버시 침해가능성 최소화 위한 제도적·기술적 장치 마련, 기업의 개인정보 관리수준 검증, ‘개인정보보호 미래 법제정비 포럼’ 운영 등
⑦ 빅데이터 산업진흥을 위한 법제도 개선	빅데이터 관련 법제도 개선 사항 발굴, 종합적인 정책방안 도출

자료: 국가정보화백서(2012) p.58, 방송통신위원회 보도자료(2012) 재구성

한국석유공사는 고유가에 따른 소비자의 부담을 줄이기 위해 유가정보를 예측·제공하는 오피넷 시스템을 구축하여 운영하고 있다. 국제 유가를 기반으로 1주일 뒤에 국내 정유사와 지역별 판매가격을 예측하여 자료로 제공하고 있으며, 국내 1,300여 개 주유소의 현재 유가를 국민에게 제공하여 인근 주유소 간에 가격 경쟁체제를 유도하고 있다. 국립보건연구원은 ‘중앙인체자원은행’을 운영하여 한국인의 질병 원인 규

10) 국가정보화전략위원회(2012), p.3

명, 진단, 치료법 연구·개발에 필요한 인체자원을 확보 및 제공하고 있다. 전국 17개 병원에서 52만 명의 인체자원 정보를 확보하였으며, 2011년까지 25만 명의 정보를 454개 보건의로 분야 연구개발(논문 159편)에 제공하였다.¹¹⁾

(2) 기업 비즈니스 분야

빅데이터의 가치가 부각되면서, 각 기업들은 보유한 데이터들을 활용하여 비즈니스 전략의 수립 및 관련 사업의 수행에 앞장서고 있다.

1) 국외 빅데이터 활용 사례

구글, 애플, 페이스북, 아마존 등 인터넷을 기반으로 핵심 서비스를 무료 또는 저가로 제공하는 기업들은 이용자들의 방대한 데이터를 수집·활용하여 비즈니스 전략을 세우고 있다.

○ 구글

구글은 하루 평균 6억 2,000만 명이 방문하고, 10억 건의 검색, 72억 건의 페이지 뷰 트래픽을 발생시킨다. 구글은 방문자의 검색어를 바탕으로 광고나 링크, 음식점 리뷰, 여행정보, 지도 데이터, 교통정보 등 일상생활과 밀접한 각종 정보들을 제공하고 있으며, 안드로이드 디바이스를 통한 사용자 정보를 보유하고 있다. 또한 검색 및 광고 서비스 운영에서 축적된 노하우를 활용하여 빅데이터 분석과 분산처리 등을 선진적으로 수행하고 있으며, 클라우드 컴퓨팅, 빅데이터 분석틀, 웹 애플리케이션 플랫폼을 통합한 빅데이터 플랫폼을 구축하여 웹 생태계의 지배적 위치를 강화하기 위해 노력하고 있다.¹²⁾

구글은 독감에 대한 인터넷 검색통계를 이용하여 28개국의 독감 유행 수준을 예측하였는데, 이를 통해 미국 질병통제예방센터(CDC)보다 1주일 이상 앞서 독감 유행을 예측하는 성과를 얻었다. 또한 미국 정부는 신차 구입의 수요가 적을 것으로 예상하고

11) 상동

12) 채승병 외 2인(2012), p.9

4개월 시행 목표로 10억 달러를 편성하였으나, 구글은 검색통계를 이용하여 오바마 정부의 노후 차량 보상 프로그램의 호응도를 분석하여 정부 예산이 부족할 것을 예측하였다. 구글의 예측대로 1주일 만에 예산이 소진됨에 따라, 미국 정부는 20억 달러를 긴급 추가하여 예산을 재편성하였다.¹³⁾

○ 페이스북

페이스북은 20억 명의 회원 및 1,000억 건의 친구관계를 토대로 하루 평균 2억 5,000만 장의 사진이 업로드되고 있으며, 27억 건의 ‘좋아요’와 댓글이 생성되고 있다. 그리고 페이스북은 회원의 관심사, 소속, 결혼 여부, 심리상태, 위치정보 등의 소셜 데이터를 보유하고 있다. 최근 페이스북은 실시간 입찰 광고 플랫폼 ‘FBX’을 통해 이용자들의 정보와 검색어를 실시간으로 분석하여 맞춤형 광고를 제작하는 등 새로운 비즈니스 전략을 구사하고 있다.¹⁴⁾

○ 애플

아이폰, 아이패드, 아이팟 등 멀티미디어 단말기를 제작·판매하는 애플은 아이튠즈 스토어, 아이클라우드 서비스를 제공하면서 이용자들의 데이터를 수집·분석하고 있다. 특히 애플의 음성인식 서비스인 ‘시리(Siri)’는 이미 구축된 데이터베이스를 토대로 이용자의 질문이나, 행동을 미리 예측해 최적의 답을 제시해 줌으로써 빅데이터의 새로운 양상을 제시하고 있다.¹⁵⁾

○ 아마존닷컴

아마존은 하루 평균 440만 명이 방문하며, 2010년 크리스마스 시즌 동안에 900만 개의 상품을 판매한 미국의 대표적인 전자상거래 사이트이다. 아마존은 1억 2,000만 명의 고객정보를 보유하고 있으며, 230만여 종의 서적 데이터를 보유하고 있다. 아마존은 회원들 개개인의 상품검색 및 구매패턴을 분석하여 자동화된 알고리즘을 통해

13) 채승병(2011), p.6

14) 《이데일리》(2012. 9. 14)

15) 《아이뉴스》(2011. 12. 24)

이용자 개인에게 맞는 쿠폰을 자동화된 알고리즘으로 생성하여 고객에게 전달한다.¹⁶⁾

○ T-mobile

미국의 T-Mobile은 3,000여만 명의 가입자의 통화 및 송수신 내역을 활용하여 고객관리시스템을 구축 및 운영한다. T-Mobile은 매일 179억 건 이상 발생하는 빅데이터를 분석하여 고객의 통신사 전환 위험을 사전에 감지하고, 소셜네트워크를 분석하여 이탈 징후를 보이는 고객에게 맞춤형 혜택을 제공하여 타사로의 통신사 이동을 절반 수준으로 방지하는 효과를 얻었다.¹⁷⁾

○ O₂-Placecast 서비스

영국의 통신사 O₂는 Placecast와 협력하여 위치기반 서비스를 이용하여 실시간으로 스타벅스 프로모션을 모바일 서비스 가입자에게 제공했다. 가입자가 스타벅스 매장 근처에 있을 경우 문자메시지로 프로모션 쿠폰을 전송하는 서비스로, 스마트폰 이용자의 50% 이상이 위치기반 서비스를 이용한 프로모션을 통해 물품을 구매한 경험을 가지고 있다.¹⁸⁾

<표 7> 기타 기업의 빅데이터 활용 사례

산 업	기 업	특 징
유 통	월마트	각 매장의 모바일과 소셜쇼핑의 특징을 이용한 '월마트랩' 운영
	자라	전 세계 매장 판매·재고 데이터 분석으로 무재고 시스템 실현
	이베이	고객의 SNS 활동내용과 과거 구매이력을 분석하여 명절·기념일에 선 물리스트 작성 및 추천
서비스	넷플릭스	고객이 대여한 영화목록 등을 분석해 개인별 맞춤형 영화 콘텐츠 제공
	디사이드닷컴	전자제품 가격흐름을 예측해 고객에게 적절한 구매시기를 알려줌
	웹포인트	환자 차트, 병원 시술자료, 논문 등 모든 정보를 검색·분석하여 적절한 치료법 제시

16) 《매일경제》 (2012. 9. 24)

17) 채승병 외 2인(2012), p.4

18) 김상락 외 2인(2012), p.21

산 업	기 업	특 징
제 조	볼보	모든 차량 내부에 센서를 부착해 이동 중 발생하는 결함 발견
	포드	자사 차량의 내부통신망에서 수집한 정보를 스마트폰에서 이용 가능한 'OpenXC' 프로젝트 추진
	히타치플랜트 테크놀로지	크레인 곳곳에 센서를 부착해 무게중심 이탈 여부, 오작동 징후를 파악하는 '크레인 닥터 클라우드' 시스템 구축
	마이크론 테크놀로지	제품 생산시간에 영향을 미치는 요소를 분석해 비용절감 방안을 수립
	코노코필립스	석유 시추선 부품 데이터 분석을 통해 장애발생 징후 패턴 발견, 필요 부품의 사전 공급 및 교체

자료: 국가정보화전략위원회(2012) p.6, 장영재(2012) p.12, 매일경제(2012. 9. 24) 재구성

2) 국내 빅데이터 활용 사례

○ 포스코

포스코는 SAS를 통해 생산 공정별 온도, 습도, 압력, 성분 등의 비철강 생산 전 과정의 각종 데이터를 0.001초 단위로 수집·분석함으로써 불량률을 최소화하여 생산 효율성이 높아지도록 실시간으로 공정제어를 한 결과, 6시그마(100만 개당 3.4개)를 달성하였다. 또한 가격변동이 큰 철광석 등의 자원을 적시에 조달하기 위해 데이터 분석을 통해 최적 구매 시기와 가격대를 결정하고 있다. 고객의 수요 데이터, 남미 및 호주 광산의 상황, 런던 금속 거래소의 광물 가격 데이터를 분석하여 미래의 철광석 가격을 예측하고 있다.¹⁹⁾

○ SK텔레콤

SK텔레콤의 티맵 내비게이션은 콜택시, 유류 운반차량, 고속버스 등에 위성위치확인시스템(GPS)을 장착하여 실시간으로 전국 도로의 교통정보를 제공하고 있다. SK텔레콤의 티맵 내비게이션은 기존의 내비게이션이 내장된 지도와 길 안내 프로그램을 활용하는 것과는 달리, SK텔레콤 서버의 슈퍼컴퓨터가 실시간 정보를 토대로 목적지까지의 최단거리를 계산하고 정확한 도착시간을 제공한다는 점에서 이용자들의 편의

19) 국가정보화전략위원회(2012), p.4

를 제고하고 있다.²⁰⁾

또한 SK텔레콤은 지도와 연결된 유동인구, 업종별·월별 매출 등의 정보(10종)로 상권분석 서비스를 제공하고 있다. 자영업 창업희망자인 이용자는 2,650만 SK텔레콤 가입자 동선(유동인구), 3,000만 OK캐쉬백 회원 소비패턴, 현대카드 가맹점 결제, 부동산 114의 상권 시세 등의 데이터를 통해 업종별 매출 현황, 경쟁 매장, 잠재 수요 고객, 유동인구 등의 정보를 지도에서 직접 분석할 수 있다.²¹⁾

○ 다음소프트

국내 소셜미디어 분석 서비스 전문기업인 다음소프트는 소셜네트워크의 이슈와 관심 키워드의 실시간 모니터링으로 상황에 맞는 대응 전략 및 마케팅 전략의 수립을 지원한다. 다음소프트에서 운영하는 ‘소셜메트릭스(socialmetrics.co.kr)’는 자연어 처리, 데이터 마이닝 등의 분석틀을 활용하여 블로그와 트위터에서 형성되는 트렌드와 여론을 일반과 공공으로 구분하여 제공하고 있다.²²⁾ 또한 ‘와인오퍼’를 개발하여 와인 가격, 어울리는 음식, 테스트, 포도품종 등 와인과 관련된 3만 5천여 개의 정보를 제공하고 있으며, 사용자 관심사 및 상황에 맞게 와인을 추천하고, 이용자들의 다양한 와인 랭킹을 제공하고 있다.²³⁾

한국은 높은 IT 인프라 구축 및 세계 최고 수준의 기술수용력을 가진 이용자를 보유한 데이터 생산국이자, 소비국이다. 하지만 데이터의 활용 측면에서는 아직도 낮은 수준에 머물러 있는 실정이다. 실질적으로 빅데이터의 가치와 효용성을 제대로 숙지하여 이를 비즈니스에 접목한 우리나라 기업은 극소수에 불과하다. 이는 ① 빅데이터가 수집되지 않는 비효율적인 비즈니스 현장, ② 역량 축적으로 이루어지지 않는 소모적인 데이터 활용 상태, ③ 빅데이터 관리와 분석에 필요한 지식기반 취약 등에 기인한다.²⁴⁾

20) 김상탁 외 2인(2012), p.21

21) 국가정보화전략위원회(2012), p.4

22) 《디지털타임즈》(2011. 10. 10)

23) 《디지털타임즈》(2012. 6. 14)

24) 채승병 외 2인(2012), pp.15~17

3. 빅데이터 관련 시사점

(1) 개인정보 및 지적재산권 보호

최근 미국 연방거래위원회(FTC)는 구글이 맞춤형 광고 제작을 목적으로 애플의 전용 웹브라우저 ‘사파리’를 사용하는 고객의 개인정보를 무단으로 수집한 것과 관련하여 벌금 2,250만 달러(한화 약 253억 원)를 부과하였다.²⁵⁾ 또한 글로벌 IT 기업인 IBM은 임직원들에게 애플의 ‘시리’ 서비스 사용 금지령을 내렸는데, 이는 직원들의 음성 검색 내용이 애플의 서버에 축적됨에 따른 사내 정보 유출을 우려했기 때문이다.²⁶⁾

이처럼 많은 기업들이 비즈니스 전략으로 서비스 이용자들의 동의와 무관하게 각종 개인정보를 수집·거래하면서 사회적인 문제로 부각되고 있으며, 빅데이터의 활용이 늘어나면서 그 피해양상은 더욱 커질 것으로 예상된다. 예컨대 우리는 주요 장소 및 거리에 설치된 CCTV가 실시간으로 보행자정보를 취합하고, 자동차 블랙박스가 실시간으로 운전정보를 데이터화하며, 스마트폰 보급 및 SNS의 활성화로 개인의 사소한 대화까지도 수집·분석되고, 각종 정보를 손쉽게 공유할 수 있는 사회에서 살고 있다. 하지만 개인정보 유출 및 지적재산권 보호에 관련된 안전망은 아직 구축되지 못하고 있다. 따라서 각종 정보가 ‘정보의 홍수’ 속에 무방비 상태로 표류하고 있는 실정을 고려하여, 개인 프라이버시 및 지적재산권을 과도하게 침해하지 않는 범위 내에서 데이터를 잘 활용할 수 있도록 하는 가이드라인의 마련이 시급하다.

(2) 데이터 분석 인프라 구축

빅데이터의 가치가 높게 평가되면서 정부와 국내 기업들이 데이터의 수집·관리에 많은 노력을 기울이고 있다. 하지만 국내 빅데이터 산업은 기술력 및 자금력의 한계로 주로 IBM, EMC, HP, SAS, Oracle 등 글로벌 기업의 기술과 오픈소스 솔루션에 의

25) 《매일경제》(2012. 8. 10)

26) 《조선비즈》(2012. 6. 22)

존하고 있다.

빅데이터의 활성화를 위해서는 대량의 데이터를 수집·분석할 수 있는 하드웨어, 데이터 플랫폼, 분석기술(자연처리, 의미분석, 데이터 마이닝)과 같은 소프트웨어가 개발되어야 하며, 수학·공학적인 능력과 경제학·통계학·심리학 등 다방면에 능통한 인재가 필요하다. 이를 위해 대용량 분산파일시스템, 시스템 가상화 기술, 기계학습 및 인공지능 분야 등의 빅데이터의 핵심 기술을 개발할 수 있도록 국가 차원에서의 마스터플랜 추진 및 지속적인 지원이 요구된다. 또한 빅데이터 처리 인프라 개발 및 분석 플랫폼 구축을 위한 ‘시스템 엔지니어’, 데이터 분석의 분명한 목적의식과 통합적 사고, 해석력을 갖춘 ‘데이터 사이언티스트’, 직관적인 통찰과 풍부한 정보를 동시에 제공하는 ‘인포그래픽스 아티스트’ 등의 핵심 인력을 양성하기 위한 인재육성 프로그램 및 교육과정의 개설이 필요하다.²⁷⁾

4. 결 어

2012년은 ‘빅데이터의 원년’으로 삼을 정도로, 전 세계적으로 빅데이터에 대한 논의가 다양하게 이루어졌다. 1월에 열린 세계경제포럼(WEF)에서는 빅데이터 기술을 국제개발의 새로운 가능성을 여는 2012년의 가장 중요한 기술로 지목했으며, 3월에 미국은 ‘빅데이터 연구개발 이니셔티브’를 발표하였고, 영국은 ‘데이터 전략위원회’를 설립하였으며, EU는 ‘데이터 개방전략’ 등의 빅데이터 정책들을 수립하였다. 우리나라도 ‘빅데이터를 활용한 스마트정부 구현방안’을 토대로 빅데이터 국가전략포럼을 실시하고, 빅데이터 마스터플랜을 추진하였다.

또한 공공·행정 분야 외의 기업 비즈니스 분야에서는 전략적으로 빅데이터를 활용하여 높은 가치를 창출하고 있다. 구글, 페이스북, 애플, 아마존 등의 인터넷을 기반으로 한 글로벌 기업들뿐 아니라, 유통, 제조 등의 다양한 사업 분야에서도 업무효율성

27) 안창원·황승구(2012), p.16

제고 및 의사결정 능력 향상, 경영상 문제 발견 및 해결에 빅데이터를 적극적으로 활용하고 있다. 반면, 국내 기업은 빅데이터 관리와 분석에 필요한 지식기반 취약과 인프라 부족 등으로 빅데이터를 경영에 접목시킨 사례가 많지 않으며, 주로 글로벌 기업의 기술과 오픈소스 솔루션에 의존하고 있다.

한편, 빅데이터의 중요성이 강조되면서 개인정보 및 지식재산권 보호, 데이터 분석 인프라 구축에 대한 이슈가 부각되고 있다. 이에 따라 개인정보 및 지식재산권을 보호하면서 빅데이터 산업을 활성화시킬 수 있는 가이드라인의 마련이 시급하며, 빅데이터 산업 인프라(하드웨어, 소프트웨어, 인력) 구축을 위한 마스터플랜 추진, 산학연 간의 지속적인 협력과 지원이 요청된다.

참고문헌

- 김상락·강만모·박상무 (2012), “빅데이터가 여는 미래의 세상”, 《산업조직연구》 6월호, pp.18~24.
- 국가정보화전략위원회 (2012), “빅데이터 마스터플랜 추진 현황 및 향후 계획”.
- 방송통신위원회 (2012), “방통위, 빅데이터 서비스 활성화 적극 나선다”, 보도자료.
- 안창원·황승구 (2012), “빅데이터 기술과 주요 이슈”, 《산업조직연구》, 6월호, pp.10~17.
- 유지연 (2012), “세계경제포럼(WEF)을 통해 본 빅데이터 논의 동향과 함의”, 《방송통신정책》, 제24권 4호, 정보통신정책연구원.
- 이강용 외 (2012), “공공분야에서의 빅데이터 활용을 위한 지식자산 구축”, 《산업조직연구》, 6월호, pp.40~46.
- 장영재 (2012), “빅데이터와 비즈니스의 새로운 패러다임”, 《DIGIECO FOCUS》, KT경제경영연구소.
- 정용찬 (2012), “빅데이터 혁명과 미디어 정책 이슈”, 《KISDI Premium Report》, 12-02, 정보통신정책연구원.

채승병 (2011), “정보홍수 속에서 금맥 찾기: 빅데이터 분석과 활용”, 《SERI 경영 노트》, 제91호, 삼성경제연구소.

채승병·안신현·전상인 (2012), “빅데이터: 산업 지각변동의 진원”, 《CEO Information》, 851호, 삼성경제연구소.

행정안전부·방송통신위원회·지식경제부 (2012), “빅데이터 시대의 국가발전전략”, 『2012 국가정보화백서』, pp.49~58.

《디지털타임즈》(2011. 10. 10), “다음소프트, 대용량 소셜미디어 분석 실시간 서비스”.
 _____ (2012. 6. 14), “다음소프트, 아이폰용 와인 앱 ‘와인오피’ 출시”.
 _____ (2012. 8. 10), “정부, 재난대응 빅데이터 본격 적용”.

《매일경제》(2011. 11. 4), “SNS 읽어라, 소비자 마음이 보인다”.
 _____ (2011. 12. 13), “EU, 공공정보 공개, 61조원 시장 만든다”.
 _____ (2012. 8. 10), “구글에 253억원 벌금, FTC 개인정보 무단사용에 부과”.
 _____ (2012. 9. 24), “빅데이터 활용 아직 초보단계, 스몰데이터부터 경험 쌓아야”.

_____ (2012. 9. 24), “글로벌 기업, 빅데이터로 노다지 캔다”.

《아이뉴스》(2011. 12. 24), “수집하고 분석해 미래를 예측하라”.

_____ (2012. 3. 29), “IDC, 2015년 빅데이터 시장 규모 169억 달러 전망”.

《이테일리》(2012. 9. 24), “페이스북, 광고시장에 뛰어든다”.

《조선비즈》(2012. 6. 22), “시리가 똑똑한 비서? 사실은 은밀한 스파이”.

IDC (2012), “Worldwide Big Data Technology and Service 2012~2015 forecast”.

각국 공공데이터 공개 사이트

미국 data.gov	캐나다 data.gc.ca
영국 data.gov.uk	싱가포르 data.gov.sg
뉴질랜드 data.govt.nz	독일 portalu.de
호주 data.gov.au	덴마크 data.digitaliser.dk

스페인 datos.gob.es

핀란드 suomi.fi

그리스 geodata.gov.gr

벨기에 data.gov.be

노르웨이 data.norge.no

네덜란드 data.overheid.nl

사우디 saudi.gov.sa

오스트리아 data.wien.gv.at

프랑스 opendata.paris.fr

대한민국 data.gov.kr