

특집

건설산업, 빅 데이터 시대를 준비하라

빅 데이터-건설산업 발전에 기여할 수 있나

강 상 혁 | 한국건설산업연구원 연구위원
shkang@cerik.re.kr

건 설업 분야에서 데이터의 활용도는 매우 저조하다. 거의 모든 의사결정 상황에서 결정적인 근거를 제공하는 것은 경험과 직감이다. 데이터는 단지 지나간 흔적일 뿐 미래를 위한 어떤 용도로도 사용되지 않는다.

하지만 최근 데이터의 위상이 급격히 높아지면서 빅 데이터(big data)라는 신조어까지 등장한 마당에 제조, 유통업 분야는 빅 데이터 기술을 앞 다퉈 적용하여 상당한 효과를 보고 있다.

본고에서는 빅 데이터란 무엇이고, 그것이 건설 생산성 증진에 기여할 수 있는지, 또한 건설산업계는 이 기술을 어떻게 받아들여야 하는지에 대하여 논의하고자 한다. 참고로 본고의 내용 가운데 일부는 「한국건설관리학회지」 2013년 12월호에 실린 “건설산업과 빅 데이터(Big Data)”의 내용을 인용·참조하였음을 밝혀둔다.

빅 데이터의 정의

빅 데이터의 일차적인 의미는 종래의 방법으로는 수집, 저장, 검색, 분석이 어려운 방대한 양의 데이터이다. 이차적으로는 그런 큰 데이터를 여러 기법을 이용해 유의미한 정보로 만들어내는 과정까지를 포함한다. 보다 구체적으로 정의하자면 빅 데이터는 4V로 통칭되는 규모(volume), 다양성(variety), 속도(velocity), 가치(value)의 측면에서 정의될 수 있다.

우선, 처리되는 데이터의 규모가 메가(mega, 100만), 기가(giga, 10억) 바이트에서 테라(tera, 1조), 페타(peta, 1,000조), 엑사(exa, 100경) 바이트로 증대되었다. 처리되는 데이터의 범위에 있어서는 엑셀과 같이 규격화된 틀에 일목요연한 표로 구성할 수 있는 정형 데이터에서 동영상, 음악, 메시지, 소셜 미디어, 게시물 등과 같은 비정형 데이터로 그 종류와 유형이 다양해졌다. 또한, 유통 활용 주기는 며칠~몇 주에서 몇

특집 건설산업, 빅 데이터 시대를 준비하라

분~몇 초 단위로 빨라졌다. 마지막으로, 빅 데이터는 가치를 창출한다. 아무리 크고 다양한 데이터라 하더라도 가치를 창출하지 못하면 빅 데이터라 할 수 없다.

빅 데이터에 대한 관심 급증

〈그림 1〉은 2004년 1월부터 2013년 10월까지 구글 검색 엔진에서 'Big Data'와 'Data Mining(데이터 마이닝)'의 검색 빈도 추이를 보여준다. 차트를 보면 2011년부터 빅 데이터의 검색 빈도가 기하급수적으로 증가했다는 사실을 알 수 있는데, 이는 최근 몇 년 사이 빅 데이터에 대한 관심이 비약적으로 증가했다는 것을 의미한다.

빅 데이터를 이해하기 위해서는 우선 데이터 마이닝을 이해해야 한다. 데이터 마이닝은 대량의 데이터 집합으로부터 유용한 상관관계를 발견해 미래에 실행 가능한 정보를 추출해내고 이를 의사결정에 이용하는 과정을 의미하며, Knowledge-Discovery in Database(KDD, 데이터베이스 속의 지식 발견)라고도 일컫는다. 광산에서 금이나 다이아몬드와 같은 가치 있는 보석을 캐낸다고 해서 마이닝(mining)이라는 용어가 붙은 것이다. 데이터 마이닝은 통계 및 수학적

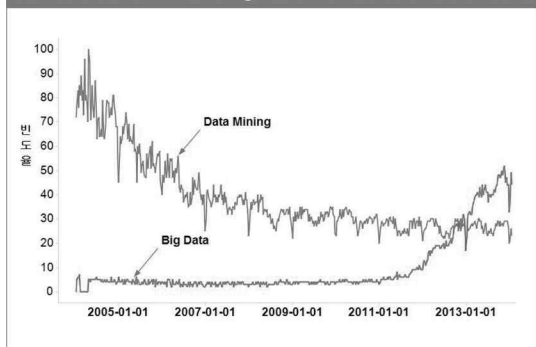
방법뿐만 아니라 다양한 패턴 인식 기술들을 이용해 데이터를 분류, 군집하고, 변수들간의 연관성을 찾아 내 미래를 예측하는 것에 가장 큰 목적이 있다.

〈그림 1〉을 보면 데이터 마이닝의 빈도는 점점 감소하고 있는데, 이것은 빅 데이터가 데이터 마이닝의 개념을 포괄적으로 섭렵하고 있기 때문이다. 기존에는 데이터 마이닝이라고 불렸으나 최근 들어 데이터의 규모가 커지고 처리 속도가 빨라지면서 빅 데이터란 용어로 대체된 것으로 사료된다.

데이터 마이닝과 함께 빅 데이터의 출현을 앞당긴 것은 스마트 기기의 보급과 데이터 처리·분석 하드웨어와 소프트웨어의 발전이다. 스마트 기기의 보급으로 개인의 행동 양식 추론을 가능케 하는 데이터가 본격적으로 양산되기 시작했고, 여기에 초대용량 데이터의 실시간 처리를 가능케 하는 하드·소프트웨어의 발달이 맞물려 '빅 데이터' 개념이 탄생하게 된 것이다.

비야흐로 전 세계적으로 빅 데이터에 대한 관심이 뜨겁다. 2012년 1월 세계경제포럼에서는 빅 데이터를 '미래의 국제 개발에 무한한 가능성을 담고 있는 기술 중 하나'로 지목했다. 2012년 3월에는 미국 백악관 과학기술정책실에서 미국의 빅 데이터 역량을 강화하기 위한 범정부적인 연구 개발 계획을 발표했다. 우리나라도 이에 발 빠르게 대처하고 있는데, 2012년 11월 대통령 소속 국가정보화전략위원회는 빅 데이터의 분석 및 활용을 실현하기 위한 비전과 미래 정책 방향을 제시했다. 또한 제4차 건설사업정보화(CALS) 기본계획(2013~2017)에서도 빅 데이터 기술 활용 저변 확대

〈그림 1〉 2004년부터 2013년까지 Big Data와 Data Mining의 검색 빈도 추이



자료 : Google Trends.¹⁾

1) Google Trends(www.google.com/trends)에서는 검색 단어(search term)에 대해 2004년부터 현재까지 주간 검색 빈도를 제공하고 있다. 여기서 제공하는 수치는 실제 검색 수가 아닌 해당 기간 최소 검색 빈도를 0으로, 최대 검색 빈도를 100으로 설정한 상대적인 값이다.

를 위한 계획을 제시하고 있다.

빅 데이터의 경제적 가치

글로벌 리서치 기관들은 빅 데이터의 활용으로 사회적 편익 증가, 생산성 향상, 에너지 소비 절감 등의 효과가 기대된다고 전망하고 있다. 경영 컨설팅 그룹인 McKinsey Global Institute(MGI)는 글로벌 비즈니스 지형을 뒤바꿀 기술 트렌드의 3가지 핵심은 '클라우드', '빅 데이터', '스마트 자산(smart assets)'이며, 빅 데이터는 혁신, 경쟁력, 생산성의 핵심 요소로 자리 매김할 것이라고 전망했다. 또한, 빅 데이터의 활용을 통해 의료, 공공 행정 등 5대 분야에서 6,000억 달러 이상의 가치를 창출할 것이고, 특히 미국의 의료 서비스 부문에서 연간 3,000억 달러의 편익과 0.7%의 생산성 증가 효과가 나타날 것이라고 예측했다. 유럽의 공공 행정 부문에서는 매년 2,500억 유로

의 편익과 0.5%의 생산성 증가 효과를 낼 것으로 기대하고 있다. 아울러 미국 소매사업 부문의 이익은 60% 증대될 수 있고, 제조업도 제품 배치와 조립 비용의 최대 50%를 절감할 수 있을 것으로 예측하고 있다.

건설 생산성 증진에 기여할 수 있는가

그렇다면 건설산업은 빅 데이터를 어떻게 받아들여야 하는가? <그림 2>는 산업 영역별 빅 데이터 활용을 통해 얻을 수 있는 잠재적 가치를 보여준다. 다시 말해 빅 데이터가 특정 산업 영역의 생산성 향상에 얼마만큼 기여할 수 있는지를 상대적으로 표현한 것이다.

MGI는 빅 데이터가 건설산업의 생산성 향상에 큰 기여를 하지 못할 것으로 예측했다. 그리고 그 이유를 네 가지 측면에서 설명하고 있다.

첫째, 빅 데이터 인력 부족이다. 빅 데이터로부터 가치를 창출할 수 있는, 통계와 기계 학습 분야 전문

〈표 1〉 빅 데이터의 경제적 가치에 대한 글로벌 리서치 기관들의 전망

기관명	주요 내용
Economist(2010)	<ul style="list-style-type: none"> • 데이터는 자본이나 노동력과 거의 동등한 레벨의 경제적 투입 자본, 비즈니스의 새로운 원자재 역할 • 비즈니스 트렌드 파악, 질병 예방, 범죄 해결 등 효과
MIT Sloan(2010)	<ul style="list-style-type: none"> • 데이터 분석을 잘 활용하는 조직일수록 차별적 경쟁력을 갖추고 높은 성과 창출 • 조직 분석 역량 3단계(열망-숙련-변혁 단계) 특징 제시
PwC(2010)	<ul style="list-style-type: none"> • 빅 데이터는 이전까지는 다루지 못하고, 시도하지 못했던 데이터의 활용을 가능하게 하며 잠재적 가치와 영향력이 높음. • 빅 데이터의 중요성에 대해 기업들이 주목하고 있으며, 새로운 비즈니스의 가치 창출에 핵심 요소가 될 것
Gartner(2011)	<ul style="list-style-type: none"> • 데이터는 21세기의 원유이며, 데이터가 미래 경쟁 우위를 좌우 • 기업은 다가올 '데이터 경제 시대'를 이해하고 정보 고립(information silo)을 경계해야 성공 가능 • 빅 데이터는 향후 주목해야 할 신흥 기술(2~5년 후 성숙)
McKinsey(2011)	<ul style="list-style-type: none"> • 글로벌 비즈니스 지형을 뒤바꿀 기술 트렌드의 3가지 핵심은 '클라우드', '빅 데이터', '스마트 자산(smart assets)' • 빅 데이터는 혁신, 경쟁력, 생산성의 핵심 요소 • 의료, 공공 행정 등 5대 분야에서 6,000억 달러 이상 가치 창출

자료 : 「신 가치 창출 엔진, 빅 데이터의 새로운 가능성과 대응 전략」, 한국정보화진흥원, 2011. 12. 30.

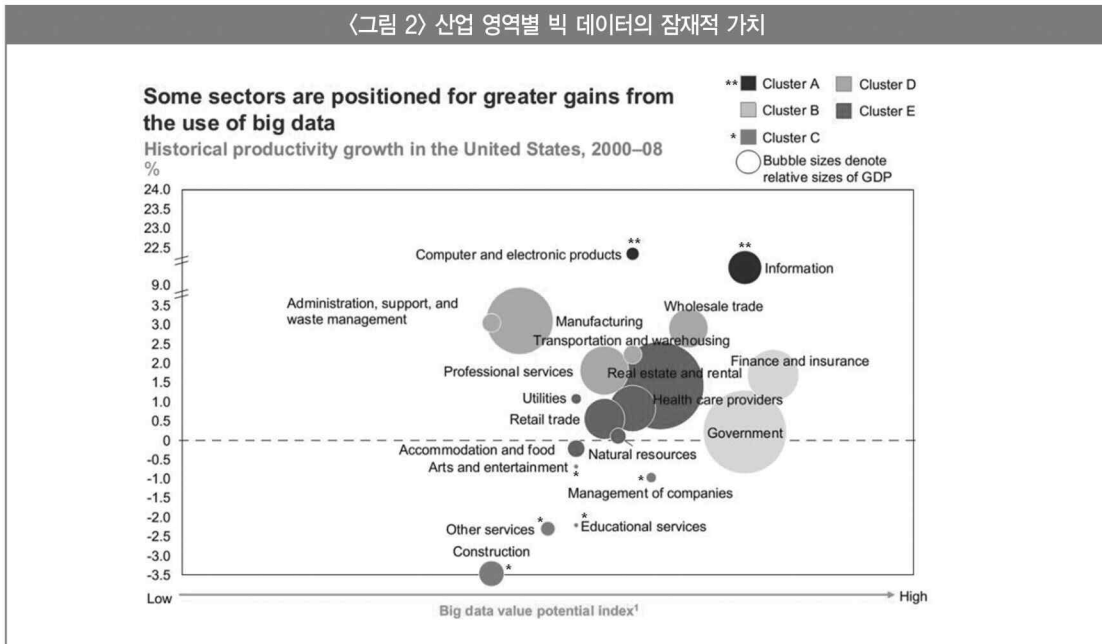
특집 **건설산업, 빅 데이터 시대를 준비하라**

지식을 가진 인력이 매우 부족하다는 것이다. 둘째, 건설 분야는 IT 집중도(intensity)가 낮다. 아무래도 빅 데이터가 디지털화된 데이터를 근간으로 하기 때문에 IT 집중도가 높은 산업 영역이 상대적으로 높은 잠재적 가치가 있는 것은 당연하다. 제조업이나 유통업과 같은 산업과 비교해볼 때 건설 영역은 IT 집중도가 낮은 것이 사실이다. 셋째, 건설 분야는 데이터 중심의 사고 습관(data-driven mind-set)이 형성되어 있지 않다. 맞는 얘기다. 데이터보다는 경험이나 관습에서 해답을 찾으려는 경향이 유독 강한 분야가 바로 건설 분야이다. 마지막으로, 데이터 유효성·입수 가능성(data availability)이다. 건설 분야는 데이터가 충분히 쌓여 있지도 않을 뿐더러 실시간으로 데이터가 축적될 수 있는 환경이 조성되어 있지 않다는 것이다. 이에 반해 제조업 및 유통업계를 보면 고객의 결제 정보가 실시간으로 중앙 데이터베이스에 기록되고, 이

를 바탕으로 고객의 취향을 읽어내는 데까지 불과 며칠밖에 걸리지 않는다고 한다. 물론 산업 영역별 고유한 특성이 있기는 하지만 빅 데이터의 관점에서 보면 건설 영역의 데이터 유효성은 저조해 보인다.

MGI의 분석은 일리가 있다. 건설 분야에 빅 데이터 인력이 부족하고, 데이터 중심의 사고 습관이 형성되어 있지 않다는 것은 인정하지 않을 수 없다. 하지만 MGI는 건설을 너무 협소한 개념으로 본 측면이 있다. 아마도 건설을 현장에서 발생하는 '건물을 세우고 도로를 포장하는 등의' 시공 위주의 개념으로만 본 듯하다. 그러나 건설은 그 스펙트럼이 넓고 세부 분야도 매우 다양하다. 시공만이 건설이 아니라 기획, 설계, 구매, 시공, 시운전, 유지관리에 이르는 생애주기가 건설의 대상이라는 뜻이다. 이렇게 확장된 개념으로 바라보면 건설산업이야말로 빅 데이터를 통해 높은 생산성 향상 효과를 볼 수 있는 분야가 아닐 수 없다.

〈그림 2〉 산업 영역별 빅 데이터의 잠재적 가치



자료 : Big data - The next frontier for innovation, competition, and productivity, p.9, McKinsey Global Institute(MGI), 2011. 5.

건설 분야의 IT 집중도는 가파른 증가 추세이다. 우리의 건설산업은 과거 문서 위주의 보고서(데이터) 생산 체계에서 이제는 PMIS와 같은 디지털화된 데이터 기반의 업무 처리 환경으로 진화하였다. 최근에는 BIM과 같은 혁명적 변화의 바람까지 불고 있으니 건설 분야의 IT 집중도는 결코 낮다고 볼 수 없다. 다만 타 산업에 비해 다소 늦을 뿐이다.

이와 더불어 디지털화된 데이터의 양산이 이루어짐에 따라 유효하고 입수 가능한 데이터는 점증하고 있다. 가까운 예로 고속도로와 같은 인프라 시설물의 운영 및 유지관리 업무를 생각해 보자. 하루에 톨게이트를 이용하는 차량만 수백만 대이고, 최근에는 전자식 단말기(하이패스)를 통해 통행 정보가 자동으로 기록되고 있다. 또 다른 예로 교량 계측 센서로부터 입수되는 방대한 용량의 데이터를 생각해 보자. 최근에는 센서 기술과 디지털 데이터의 저장 기술이 발달하여 교량의 변위, 진동, 변형, 기울기, 크랙과 같은 데이터가 실시간으로 기록되고 있다. 또한 공사 현장에서 무수히 쏟아지는 업무 보고서와 같은 데이터는 조금만 관점을 달리해서 보면 '다이아몬드를 품고 있는 원석'이 아닐 수 없다.

문제는 이러한 데이터로부터 어떻게 인사이트(Insight)와 포어사이트(foresight)를 추출해 내느냐 하는 것이다. 반갑게도 세계 곳곳에서 이러한 시도가 목격되고 있다. 최근 네덜란드에서는 재해 예방 시설물 모니터링과 관리를 위해 빅 데이터 기술을 적용하는 작업에 착수하였다고 한다. 호주의 한 업체는 광산이나 건설 현장의 안전사고를 예방하기 위해 현장에서 생산되는 데이터를 실시간으로 분석하여 작업자들에게 안전에 대한 조기 경고를 알리는 시스템을 개발하여 현장에 적용하고 있다. 호주의 멜버른에서는 트

램(tram) 시스템을 빅 데이터 체계로 전환하여 도시에서 급작스레 발생하는 이벤트에 대해 실시간 대처가 가능하도록 하였다. 멜버른 교통 관계자들은 빅 데이터 기반의 교통 시스템이 호주 멜버른이 3년 연속 세계에서 가장 살 만한 도시로 선정되는 데 일조했다고 자부한다.

미래 생존의 핵심 자산

데이터 자체는 아무것도 아니다. 그것은 우리에게 어떠한 정보도 제공해주지 못한다. 그러나 데이터에서 패턴이나 연관성을 찾아낼 때 우리는 상당한 통찰(Insight)을 얻을 수 있다.

건설업계 종사자들이 경험과 관습을 토대로 사고하고 행동하는 것을 좋아한다는 사실은 인정한다. 하지만 그와 같은 직감도 중요하지만 수학적 사고도 반드시 필요하다. 왜냐하면 수학적 사고는 직감과는 별개의 방식으로 정보를 분석해 현상을 정확히 이해하도록 돕기 때문이다. 이것이 빅 데이터의 존재 가치이고 우리에게 필요한 이유인 것이다.

건설산업의 심연에는 관리 시스템이나 분석 도구에 대한 저항감 내지 반감이 존재한다. 기업체들은 오랜 기간 동안 엑셀에 의존해 왔고, 프로젝트 관리 소프트웨어는 아직도 현장에서 사용하기를 꺼려하고 있으며 사실상 활용조차 되지 않고 있다. 하지만 빅 데이터는 미래의 경쟁 환경 속에서 생존을 위한 핵심 자산이 될 것이라는 전문가들의 의견을 좌시해서는 안 될 것이다. 건설업계는 데이터의 중요성과 잠재 가치가 미래로 갈수록 증대될 것이라는 사실을 깊이 인식하고, 데이터의 활용 역량 수준을 단계적으로 제고하는 계획을 세워야 할 것이다. CERIK