

# 빅데이터를 활용한 부동산시장 분석 및 예측모형 개발



정 건 섭 교수  
부경대학교 행정학과



김성우 선임연구원  
부경대학교 공공정책연구소

## I. 서론

저출산 고령화의 영향으로 주택소비의 핵심계층인 30-54세의 비중이 2010년 49.2%에서 2020년에는 39.4%로 감소하고, 2035년에는 29.3%로 줄어들 것으로 추계되고 있다. 반면, 65세이상의 노인인구 비중은 2010년 17.8%에서 2020년에는 24.0%, 2035년에는 40.5%로 급격히 증가할 것으로 예측하고 있다. 한편, 1인가구의 비중은 노인인구 증가보다는 완만한 2010년 23.9%에서 2020년 30.4%, 2035년 34.3%로 예상된다.

특히, 부동산자산의 비중이 80%전후인 베이비부머의 은퇴와 자녀의 분가, 배우자의 사망에 따른 가구원 수의 감소 현상을 보이고 있는 고령가구의 증가는 주거공간의 규모뿐만 아니라 주택의 유형, 점유형태 등 주택소비의 조정을 가져올 것으로 예상된다. 따라서 이러한 사회·인구구조의 변화에 대응하는 맞춤형 부동산 정책에 대한 요구가 날로 증가하고 있다.

전통적인 주택의 임차구조였던 전세시장이 붕괴되고 있고, 하우스 푸어(house poor), 렌트 푸어(rent poor)라는 신종어가 등장하였다.

하우스 푸어(house poor)는 주택가격이 상승할 것이라는 기대감으로 무리한 대출로 주택을 구입하였으나 부동산 가치하락으로 인해 금융대출 상환에 어려움을 겪고 있는 경우로 사회적 잠재적인 위험요소이다. 하지만 하우스 푸어에 대한 규모의 추정은 기관마다 차이를 보이고 있어 막연한 상태에서 정책 수립이 이루어지고 있는

실정이라 해도 지나친 과언은 아닐 것이다.

〈표 1〉 하우스 푸어(house poor)의 규모

| 기관                   | 규모                        |
|----------------------|---------------------------|
| 한국경제연구원              | 광의 156.6만가구, 협의 108.4만 가구 |
| KB 금융경영연구소           | 81.0만가구(강통주택 18.5만 가구)    |
| 금융위원회/<br>기획재정부/한국은행 | 잠재적 위험 56.8만 가구           |
| 한국금융연구원              | 잠재적 위험 47.6만 가구           |
| LG경제연구원              | 32만 가구                    |

렌트 푸어(rent poor)는 최근 전세가격이 급등하면서 전세자금 마련을 위해 어려움을 겪고 있는 전세빈곤층을 의미하는 것으로 저금리 상황에서 우리나라의 전통적인 임대차 방식인 전세가 감소하고, 보증부 월세 또는 월세형태로 주택 임대시장의 구조적 변화에 기인한다. 가계 자산 중에서 부동산을 포함한 실물자산의 비중이 높은 우리나라의 경우 부동산 가격의 불안은 가계자산 안전을 위협하는 주된 요인이 되고 있다.

부동산시장에 금융기법이 도입되면서 부동산과 금융시장의 상호의존성이 높아지고 있다. 세계화에 따른 글로벌 금융체제의 구축은 부동산 시장의 대외적 의존도를 높이고 있다. 따라서 경제정책의 효율성과 주거복지 및 부동산시장 안정 등이 복합적으로 고려된 선제적이고

지속가능한 부동산 정책의 수단이 요구되고 있다. 복잡적이고 가변적인 부동산시장에 능동적으로 대처하기 위해 부동산시장에 대한 다양한 정보의 수집 및 관리 인프라 마련과 더불어 분석, 예측 등을 통해 과학적이고 합리적인 정책결정 지원체계 마련이 필요하다.

산업사회에서 지식정보화 사회로 변모해 가면서 사회 변화의 속도가 증가하고 있다. 사회·경제·문화 분야에서 빠르게 진척되고 있는 세계화와 삶의 질을 중시하고 가치를 추구하는 등 국민 삶의 변화와 맞물려 IT 기술변화는 고도의 참여와 공유 개방 등 전통적 사회와 다른 사회적 요구를 동시에 충족시키고 있다. 더불어 IT기술의 발전과 함께 정보, 통신, 전자 기술의 융합과 사회 구성원 간 가치의 네트워크 결합은 상호간의 의존성 및 복잡성을 강화시키고 있다. 이러한 사회의 역동성으로 인한 요소들 간의 관계를 과학적으로 이해하기 위해 다양한 데이터 기반의 융·복합형 접근이 필요하다고 하겠다.

최근 스마트폰과 SNS(social networking service)의 보급으로 실시간 소통이 이루어지면서 광대한 규모(volume), 속도(velocity), 다양성(variability)의 특성을 지닌 '빅데이터(big data)'가 발생하고 있다. IT산업의 발전으로 인해 방대한 양의 빅데이터를 처리·분석 할 수 있는 기술개발 및 여건이 마련되면서 과거에 버려지거나 방치되었던 데이터들이 새롭게 주목받고 있다. 맥킨지(McKinsey)는 빅데이터를 향후 비즈니스 지형을 변화시킬 10대 기술 트렌드 중 하나로 선정하고, 빅데이터를 수집, 저장하여 이를 토대로 정보를 마이닝(mining) 하는 것이 중요한 가치창출효과를 가져 올 것이라고 전망하고 있다.

우리 삶에서 생산되는 많은 정형 혹은 비정형 데이터는 공간과 관련 되어 있다. 공공에서 생성되고 있는 빅데이터의 80%가량이 공간정보를 기반하고 있다. 지리적(또는 공간적) 위치와 관련된 기술(또는 속성) 데이터와 융합된 공간자료를 다양한 유형의 분석에 활용함으로써 그 성과가 최종수요자에게 도달토록 하는 맞춤형 부동산 정책을 수립·집행하는데 효용성이 매우 크다고 본다. 최근 정부는 향후 빅데이터가 미래 국가경쟁력의 주요한 요소가 될 것으로 인식하고 공공정책 분야에도 빅데이터를 활용

한 사회 현안 및 행정 수요를 파악해 공공서비스의 대응성을 높일 수 있을 것으로 예상하고 있다.

하지만 빅데이터를 기반으로 한 부동산 시장 참여자의 행동을 이해하고 나아가 문제 해결과 바람직한 결과를 유도하기 위한 정책 자료로 활용되기 위해서는 부동산 빅데이터를 활용한 분석 및 예측모형의 개발과 해결해야 할 몇 가지 과제들이 있다. 이러한 맥락에서 부동산 관련 빅데이터를 활용한 정책적 방향과 과제에 대하여 언급하고자 한다.

이를 위하여 다음의 II장에서는 빅데이터의 개념 및 현황을 간략하게 살펴보고, III장에서는 빅데이터의 부동산정책 활용 방안에 대하여 언급하고, IV장에서는 빅데이터를 활용한 부동산시장 분석 및 예측모형에 대하여 설명하고, 마지막으로 V장에서 요약 및 결론을 맺고자 한다.

## II. 빅데이터 개념 및 현황

### 1. 빅데이터의 개요

빅데이터에 대한 정의는 연구자들마다 다양하게 정의하고 있다. 일반적으로 빅데이터란 기존의 데이터베이스 시스템으로 처리할 수 없는 용량을 넘어선 데이터를 말하는데(오라일리미디어, 2013:14), Newman(2011)는 이러한 빅데이터에 대한 특징을 규모(volume), 다양성(variety), 속도(velocity) 3가지로 규정하였다. IBM은 이러한 3V 요소 중 두 가지 이상을 충족시킬 경우 빅데이터 기술이라고 정의하고 있다. 빅데이터의 가장 큰 특징은 텍스트와 이미지뿐만 아니라 동영상 등 비정형성을 갖고 있다는 것이다. 규모뿐만 아니라 빠르게 전파되기 때문에 중요한 패턴을 찾기가 쉽지 않고, 유용한 정보의 증가만큼 불필요한 정보도 급증하고 있다고 한다(최진원·김이연, 2012).

데이터의 처리측면에서 Gantz & Reinsel(2011)은 빅데이터 기술을 아주 방대한 양의 다양한 데이터를 매

우 빨리 모니터링 함으로써 그냥 지나칠 수 있던 것을 발견 또는 분석함으로써 그 안에 존재하는 가치를 경제적으로 추출할 수 있게끔 디자인 되어있는 차세대 기술이며 구조라고 묘사하고 있다(이정미, 2013: 58). 안창원·황승규(2012)는 빅데이터는 데이터의 규모가 방대하고(volume), 데이터의 종류가 다양하며(variety), 데이터 처리 및 분석을 적시에 해결해야 하는(velocity) 특성을 가지고 있으며, 그 결과로 새로운 가치를 창출해 낼 수 있어야 하고, 일반적인 데이터베이스로 저장, 관리, 분석 할 수 있는 한계를 넘어서며, 기업정보, 웹, 이미지/동영상, SNS, 센서 스트림 등 정형/비정형 데이터를 모두 포함하고, 분석과 예측에 있어서 실시간 처리 등 적시성을 요구한다고 정의하고 있다.

한편 Gartner(2011)는 규모, 속도, 다양성의 특성에 복잡성(complexity)을 추가하여 4가지 특성을 제시하였다. 데이터의 복잡성이란 구조나 데이터의 획득과 처리에 드는 속도, 도메인이나 규칙, 저장타입 등 데이터의 발생과 처리, 정제 등 모든 과정이 복잡해지는 것을 말한다.

이상을 정리하면, 빅데이터는 단순한 양이 많은 것을 의미하는 개념을 넘어 규모(volume), 다양성(variety), 속도(velocity), 복잡성(complexity)의 4가지 특징을 가지고 있는 정형 또는 비정형화된 것으로 이것을 처리·분석하여 그 결과 새로운 가치를 창출해낼 수 있는 기술로 초고속 데이터의 수집, 처리, 분석, 활용이 포함된 기술로 정의될 수 있다.

## 2. 빅데이터 시장규모

빅데이터 기술은 2012년 세계경제 포럼에서 10대 기술 중 첫 번째로 선정될 만큼 과학기술적 측면뿐만 아니라 정치, 경제, 사회, 문화 등 다양한 분야에 가치 있는 서비스를 제공할 것으로 기대를 모으고 있다. 오늘날 11억 인구가 SNS를 이용하고, 이중 2억5천만명이 매일 페이스북에 사진을 업로드하고 있다(김동완, 2013: 40). 지난 10년간의 데이터의 생산량보다 앞으로 2년간 생성될 데이터양이 더 많을 것으로 예측하고 있다. 2011년 한 해 동안 생성된 전 세계 디지털 양은 약 1.8ZB(1.8조 GB)에

이르는데 두 시간짜리 HD급 영화 2,000편과 맞먹는 수치이다(김상락·강만모, 2014). 2020년의 경우 2011년에 비해 약 50배가량이 증가할 것으로 예상 된다(김민수, 2014). 데이터의 형식면에서도 과거 정형 또는 텍스트 중심의 자료였다면 최근에는 그림, 동영상, 음성 위주의 비정형 데이터가 급속도로 증가하고 있다.

Manyika et al.(2011)은 빅데이터를 활용한 주요산업의 가치는 22,3조 달러에 달하고, 미국의 경우 2018년까지 14~19만 명의 전문 인력과 150만 명의 데이터 관리 인력이 필요한 것으로 추정하고 있다. 한국정보화진흥원(2012a)은 한국의 경우 향후 5년간 약 52만 개의 한국의 빅데이터 기반 일자리가 창출될 것으로 전망하고 있다. IDC(International Data Corporation, 2012)는 전 세계 빅데이터 시장이 매년 39.4% 성장하여 2015년에 169억 달러 규모로 성장할 것으로 예측하고 있다. 국내빅데이터 시장의 경우 부동산, 교통, 교육 등 다양한 분야에서 빅데이터 서비스 발굴을 통한 성장이 예상된다. 한국정보화진흥원(2013)은 국내 빅데이터 시장이 2015년 약 263백만 달러(한화 3천억 원)에 이를 것으로 예상하고 있으며 이는 세계 시장의 약 1.6%를 점유할 것이며, 2020년에는 약 900백만 달러(한화 약 1조 원)에 이를 것으로 전망하고 있다.

## III. 빅데이터와 부동산 정책 활용

### 1. 국내외 부동산 정보화와 빅데이터 정책

미국의 부동산 시장 관련 정보는 미국은 통계청, 연방주택금융기관, 민간기관, 주택도시개발부(HUD; Department of Housing and Urban Development) 등에서 부동산 통계 생산 및 관리하고 있다. 미국 주택도시개발부의 정책연구소(Office of Policy Development and Research)에서 1978년부터 HUD USER를 구축하여 연방 차원의 주택 관련 정보들을 수집 관리하고 있다. 미국 정부는 2012년 3월, 「빅데이터 R&D 계획(Big Data R&D

Initiative)을 발표하며 빅데이터를 활용한 투명하고, 효율적이며 혁신적인 정부서비스 제공을 추진하고 있다. '오픈 데이터 포털 사이트(www.data.gov)'를 통해 정형, 비정형 데이터를 이용하기 쉬운 형태로 개방하고 있다.

일본에서는 총무성 통계국, 국토교통성, 일본 부동산연구회, 부동산유통기구, 도쿄증권거래소 등에서 부동산 관련 통계 생산 및 관리되고 있다. 가격과 거래와 관련된 부동산 통계는 일본 부동산 연구소와 부동산 유통기구에서 담당하고, 부동산 공급과 현황은 총무성 통계국, 국토교통성에서 관리되고 있다. 일본 국토교통성은 부동산거래시스템 개선 방안과 REINS와 연계한 Real Estate Information Center(가칭) 구상을 밝히고 있다(Kobayashi, 2013). 부동산 정보센터(가칭)는 REINS와 연계하여 거래 정보, 공공정보, 지역 및 인구정보를 통합하여 종합 부동산 정보 제공시스템이다. 거래 정보는 가격, 소유자변경이력, 담보, 건축비, 감정평가, 진단 보고서 등이다. 공공정보는 세금, 용도지역지구, 침수이력, 유지관리이력, 부동산 지도 등이고, 기타 인구, 지역 정보 등을 제공한다. 2012년 ICT에 기반한 빅데이터 관련 국가 주도의 종합적 진흥정책인 'Active Japan ICT'계획을 통해 빅데이터를 실시간으로 수집·전송·해석하여 재난관리를 포함한 정책 과제해결을 위한 제도적 틀을 마련하였다.

국내의 경우 35종 부동산통계를 국토교통부, 통계청, 한국은행 등 7개 기관이 개별적으로 생산 및 관리되고 있다. 이러한 분산적 관리로 인해 작성 내용의 중복, 공표시기의 불일치, 통계 간 연계활용 부족 등의 다양한 문제가 발생하고 있어 이를 극복하기 위한 일환으로 "부동산행정정보화일원화사업"이 추진 중에 있다. 부동산 빅데이터 관련해서는 2013년부터 공간 빅데이터 체계 구축을 위한 전략을 수립하고 기반구축사업, 공간정보 오픈 플랫폼 고도화 사업 등을 추진하고 있다.

## 2. 빅데이터의 부동산 정책 활용

빅데이터의 경우 상당부분 위치와 시간이라는 정보를 갖고 있다. 예를 들어 독감증세에 대하여 검색하거나, 부동산 거래를 위해 부동산 정보를 검색할 경우 해당 정보의 위치와 검색

시간 등이 여기에 포함된다. 빅데이터는 시간과 장소 혹은 위치정보를 기반으로 다양한 빅데이터와 융합되면서 활용 가치가 증가한다. 실세계에 존재하는 모든 정보의 80% 이상이 위치 또는 공간과 관련되어 있는 상황에서 공간정보는 빅데이터 분석에서 결코 빠뜨릴 수 없는 정보라고 할 수 있다(김민수, 2014). 특히 위치와 같은 정보는 시간이 지남에 따라 매우 빈번하고 다양하게 발생하는 대용량의 특성을 갖고 있다.

시간과 공간 정보를 포함한 빅데이터는 언제 어느 지역에서 어떤 현상이 일어나는지를 파악할 수 있기 때문에 효과적으로 사회환경에 대응할 수 있다. 따라서 물리적 환경에 대한 분석뿐만 아니라 사회·경제적 환경·인간의 행태 분석에까지 공간을 융·복합함으로써 복잡한 사회 문제 해결에 도움이 될 수 있을 것이다(이영주 외, 2014).

〈표 2〉 부동산 정책과정별 빅데이터 활용 방안

| 정책과정    | 기존 방법                    | 빅데이터 활용 방법                | 차이점                      |
|---------|--------------------------|---------------------------|--------------------------|
| 현과과제 진단 | · 통계자료 분석<br>· 설문 등 조사   | · SNS 등 실시간 자료수집 및 분석     | · 대용량의 자료<br>· 비용과 시간 절약 |
| 대안마련    | · 전문가 집단                 | · 계획지원체계 활용<br>· 실시간 의견수렴 | · 계획과정 투명화<br>· 다양한 의견수렴 |
| 의사결정    | · 전문가 집단                 | · 대안별 파급효과 시뮬레이션          | · 과학적이고 투명한 의사결정         |
| 집행      | · 집행과정에 대한 설문조사 등        | · 실시간 의견수렴<br>· 실시간 효과분석  | · 집행에 따른 반응과 효과 실시간 확인   |
| 평가      | · 통계 자료로 확인<br>· 설문 등 조사 | · 실시간 의견수렴<br>· 실시간 효과분석  | · 집행에 따른 반응과 효과 실시간 확인   |

자료 : 김대중(2014: 45)

부동산 문제 해결을 위한 정책결정과정별로 빅데이터는 〈표 2〉와 같이 활용될 수 있다. SNS와 같은 소셜빅데이터



의 경우 현실 부동산 문제를 해결하는 아이디어가 담겨있거나, 정책 집행과정에서 실시간 의견수렴을 통해 정책집행과정에서 효율성을 높일 수 있을 것이다. 나아가 정책의 효과분석에도 활용될 수 있다. 이들 빅데이터에 대한 축적과 분석기법의 개발은 정책과급효과에 대한 시뮬레이션 기능을 강화해 의사결정과정에서 정책대안을 분석하는데 도움을 줄 수 있을 것이다. 미국, 일본, 영국 등은 국가차원에서 빅데이터 전략을 수립하고 막대한 R&D투자를 추진하고 있다. 우리나라에서도 2011년 빅데이터를 활용한 스마트 정부를 구현하기 위한 방안을 모색하고 있다(김희수, 2014). 한편, 「공공데이터 제공 및 이용 활성화에 관한 법률」이 2013년 7월에 제정되면서 빅데이터 지원 및 활용과 관련된 제도적 장치의 기틀이 마련되었다.

국토교통부(2014)는 국가공간정보정책에 따라 2014년에는 385개의 공간정보사업을 시행하고 있다. 세부적으로 공간정보 융·복합사업, 실내공간정보사업, 부동산 행정정보 일원화 및 공간정보활용 확산사업, 공간정보기반조성사업 등에 약 2,946억 원이 투자되었다. 한편, 공간빅데이터 체계를 구축하기 위하여 2016년까지 공간정보인 지형 및 지적과 함께 행정정보인 인구, 소득, 기상정보, 부동산 가격 및 민간 정보인 SNS 및 검색정보 등에 대한 DB구축 및 분석모델 개발을 위한 사업이 단계별로 추진 될 계획이다.

## IV. 빅데이터를 활용한 부동산시장 분석 및 예측모형

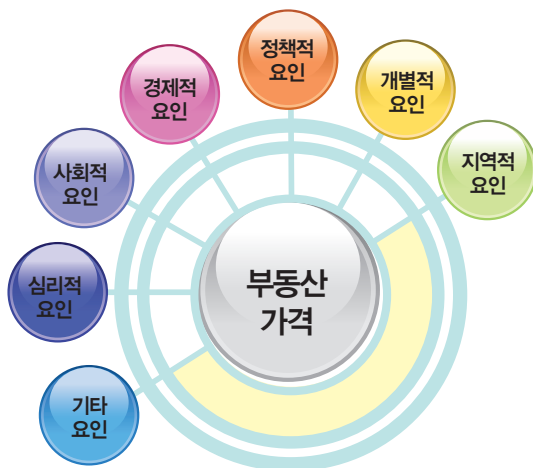
### 1. 부동산 시장예측 분석을 위한 모형 개발

부동산시장분석이란 부동산에 대한 수요와 공급에 의해 가격이 결정되는 시장 기구를 분석하는 것으로, 부동산과 관련된 정책 및 의사결정에 필요한 정보를 제공하는 일련의 과정으로 볼 수 있다. 부동산은 토지 및 정착물로 보는 법률적 시각, 소비재 또는 생산요소로서 인식하는 경제적 시각, 그리고 기술적 시각까지도 고려하는 복합적 개념을 가지고 있다. 대부분의 일반 국민과 언론은 토지와

주택 혹은 주택이라는 제한된 개념으로 부동산을 정의하고 있는 실정이다. 부동산시장에서는 개별 상품의 거래가 아닌 소유권과 같은 다양한 권리의 교환에 따라 공간의 배분과 가격이 결정된다. 또한 수급조절이 어렵고, 탄력적 공급에 따른 가격왜곡(distortion)과 부동산에 따른 지역별 시장이 형성되고, 자본시장으로의 성격도 동시에 가지기 때문에 거시경제시장과 밀접한 관련성을 가지고 있다. 공공부문에서의 부동산시장 분석은 정부의 시장 개입 시기와 강도, 그리고 정책 결과를 평가하기 위한 수단으로 활용될 수 있다. 민간 부문에서는 부동산의 취득 및 처분·개발, 투자 등을 위해 부동산시장을 분석하기도 한다.

부동산시장분석에서 중요한 범주는 수요와 공급을 형성하는 요인과 이를 변동시키는 사회, 경제, 정책 등 다양한 요소에 대한 조사 분석이다. 부동산 수요에 영향을 주는 인구 및 가구구조, 교육 및 교통, 경제상황, 산업구조, 금리, 기후 및 글로벌 시장의 변화에 대한 조사 분석이 요구됨은 물론 지역적인 성격과 외부성(externalities)으로 인해 공간적 군집을 갖는 하위시장(sub-market)이 형성되기 때문에 부동산이 위치한 지역에 대한 입지와 지역시장에 대한 조사 분석도 수반되어야 한다. 또한 현재 및 미래의 수요와 공급에 대한 요인 분석이 필요하고, 부동산이 투자재로 인식되고 있는 만큼 투자자의 심리적 요인에 대한 분석도 함께 병행할 필요가 있다.

〈그림 1〉 부동산 가격 형성요인



부동산시장과 관련된 분석 및 예측모형은 자료의 특성과 목적에 따라 거시계량모형, 부동산 관련 정책효과 분석 모형, 헤도닉가격모형(Hedonic Price Model), 공간분석모형 등이 활용되고 있다. 거시경제이론에 바탕을 둔 거시계량모형은 경제변수간의 동태적 관계를 근거로 자료의 특성에 따라 시계열모형, 회귀모형, 패널모형 등 다양한 형태로 구분된다. 한편, 헤도닉가격모형에 기초한 회귀분석 모형과 확률적 선형모형에 기초한 logit 모형, probit모형 등이 있으며, 모수의 추정방식에 따라 OLS(Ordinary Least Squares), GMM(General Method of Moment), ML(Maximum Likelihood) 추정 등이 대표적이다.

중·장기 주택시장 변동은 물론 단기 시장변동을 조기에 진단 및 예측하기 위한 구조모형 개발이 필요하다. 2000년 이후 주택시장 구조변화와 대외경제 여건변화로 인한 시장의 단기변동을 동시에 경험한 바 있으며, 이 같은 변동성은 향후 더 심해질 것으로 예상된다. 따라서 주택 정보자료의 지속적 확장, SNS 활동의 폭발적 증가, 마이크로트렌드 변화 등을 활용한 구조방정식모형을 구축하여 효과적인 주택정책의사결정이 이루어지도록 되어야 한다.

시계열분석(time series analysis)은 과거로부터 현재까지 시간의 경과에 따른 자료를 기초로 어떤 변수의 변화경향을 분석하여 미래 사회 상태를 예측하려는 동태적 분석 방법이다(서수복 2006:26). 이러한 예측 기법에는 ARIMA모형, Granger인과모형, VAR(Vector Autoregression)모형, VEC(vector Error Correction)모형, ARCH(autoregressive conditional heteroskedasticity)모형, Generalized ARCH모형 등이 대표적이다.

횡단면자료가 시계열적으로 형성된 경우 즉, 여러 개체에 대하여 그 현상이나 특성을 일련의 관측시점별로 기록해 놓은 패널데이터(panel data)인 경우, 오차항에 대한 가정에 따라 고정효과모형과 확률효과모형, 동적패널모형 등으로 구분할 수 있다. 분석 자료가 공간적인 정보를 포함하고 있는 경우에는 공간적 자기상관(spatial autocorrelation)을 가정한 공간계량 모형(Spatial

Econometric Model)이 활용된다. 대표적인 공간계량 모형은 일반공간모형(General Spatial Models), 공간자기회귀모형(Spatial Autoregression), 공간오차모형(Spatial Error Model), GWR(Geographically Weighted Regression), 베이저안 공간계량모형(Bayesian Spatial Econometrics Model) 등이 있다. 한편, GIS기법을 이용한 지구통계분석(Geostatistics)기법 등도 적용이 가능하다.

이밖에도 빅데이터를 이용한 전월세주택시장 이례현상의 지역이동을 예측할 수 있는 모형을 개발하여 주택공급 제약 및 전월세가격 급등으로 발생하는 임대차시장의 국지적 불안을 조기 예측할 수 있도록 소위 “주택임대차시장 이례현상의 공간이동 예측 모형” 개발 등이 시급히 필요하다고 판단된다. 모형 개발을 통해 전월세주택 거주 서민층의 주거안정을 제고하고, 주거비 부담이 완화될 수 있도록 함으로써 합리적 소비와 저축이라는 서민 경제의 선순환 구조를 정착하는데 기여할 수 있을 것이다. 그리고 빅데이터 분석환경에서 구현 가능한 알고리즘의 개발 및 적용의 차원에서 복수의 부동산모형 통합을 위하여 BMA(Bayesian Model Averaging) 멀티모델 앙상블의 확률적 예측과 불확실성의 정량화 및 저감 기술 개발이 필요하다.

또한 시스템 다이내믹스를 활용한 주택시장 예측시물레이션 모델 개발을 통하여 주택시장의 과학적인 진단 및 예측이 가능한 모형 개발이 이루어져야 될 것이다. 주택시장 정보공개 및 정책시행의 설명성 제고를 위한 행정업무지원 프레임워크 등은 물론 기타 시스템의 동태적 형태변화 즉, 시간의 경과에 따른 시스템의 형태변화에 초점을 둔 시스템 다이내믹스에 기초한 부동산 정책 시물레이션 모형도 정책개발 및 대안의 선택을 위해 활용되고 있다.

따라서 주택시장 변동의 원인진단 및 그에 따른 적절한 모형설계, 대단위 부동산 관련 자료를 이용한 시공간 주택시장 변동 요인의 파악, 횡단면 및 시계열 주택시장 주요 변수들에 대한 자료의 구축, 다양한 모형을 통한 주택시장 변동의 원인 진단 및 파급효과 파악 등, 이를 토대로 한 합리적인 주택정책 수립 및 시공간적 최적화 모형설계 및 검정 작업이 시급하다고 하겠다.

〈표 3〉 자료특성에 따른 부동산시장 분석 예측 모형

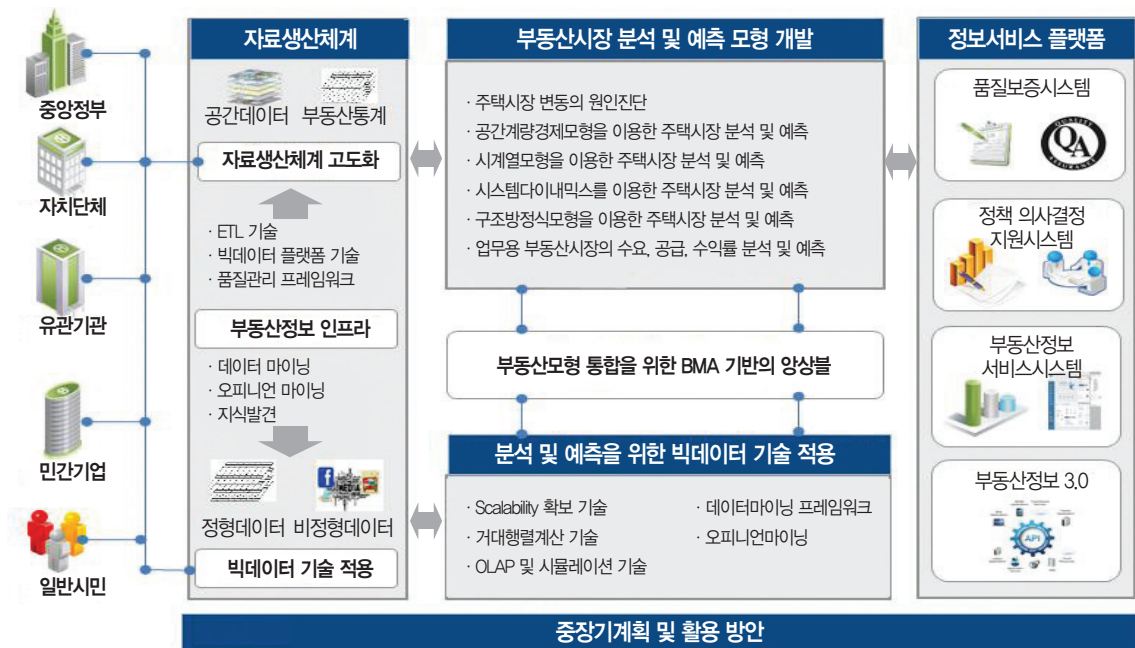
| 자료의 유형                           | 분석 및 예측 모형                                                       |
|----------------------------------|------------------------------------------------------------------|
| 횡단면 자료<br>(cross-sectional data) | 회귀분석모형, Logit모형, Probit모형, 위계적 선택모형, 신경망모형, 연립방정식모형, 구조모형(SEM) 등 |
| 시계열자료<br>(time series data)      | ARIMA모형, Granger인과모형, VAR모형, VEC모형, ARCH모형, Generalized ARCH모형 등 |
| 패널자료(panel data)                 | 고정효과모형과 확률효과모형, 동적패널모형 등                                         |
| 공간자료 (spatial data)              | 일반공간모형, 공간자기회귀모형, GWR, 베이지안 공간계량모형, 지구통계분석 등                     |
| 정책시뮬레이션                          | 시스템 다이내믹스                                                        |

## 2. 부동산시장 빅데이터 분석 시스템

부동산 빅데이터 분석 시스템이 부동산 정책 자료로 활용되기 위해서는 기존의 개별적 모형을 통해 분석 및 예측된 결과들의 통합과 최적화 추세 시뮬레이션의 수행이 필요하다. 나아가 이를 시각화함으로써 부동산시장의 수요, 공급, 가격 등과 관련한 고부가가치 정보를 탐색

및 산출 분석 그리고 예측된 결과에 대한 불확실성 관리와 감소를 위한 알고리즘이 요구된다. 특히 부동산 소셜 미디어, 뉴스 등 비정형 데이터의 수집 및 패턴의 실시간 모니터링과 분석 자료가 실시간으로 반영되어야 한다. 빅데이터를 접목하여 유용한 정보를 추출 가공하여 부동산시장을 실시간 모니터링 함으로써 선제적 대응체계 마

〈그림 2〉 빅데이터를 활용한 부동산시장 분석 및 정책서비스 개념도



련을 위해서는 세분화된 하위시장별로 국지적 불안을 조기 예측 및 이례현상을 감지하고, 부동산시장 안정화와 정책에 대한 효과를 주기적으로 모니터링 할 수 있는 능동적 대응체계마련이 필요하다. 다음의 <그림 2>는 부동산시장 분석 및 정책서비스 개념도를 나타낸 것이다.

이러한 빅데이터를 활용한 부동산시장 분석 및 예측모형을 통하여 (i) 빅데이터 기반 부동산 자료생성체계 고도화, (ii) 빅데이터 기반 부동산 시장 분석 및 예측모형 개발, 그리고 (iii) 빅데이터 기반 부동산시장 정부서비스 플랫폼 등이 구축 될 것으로 예상된다.

<표 4> 빅데이터 기반 부동산 자료생성체계 고도화

| 현재(As-Is): 부동산통계의 생산 및 관리의 문제로 인해 국가통계 신뢰성 저하와 정책 활용의 혼신 유발                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 미래(To-Be): 빅데이터 기술 기반 부동산 자료생성체계 고도화                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>▶ <b>부동산 통계생성</b></p> <ul style="list-style-type: none"> <li>· 국내 부동산 관련통계는 5개 분야 35종으로 이중 보고통계 21종, 가공통계 4종, 조사통계 10종</li> <li>· 부동산 통계를 생성 및 관리는 7개 공공기관과 2개의 민간기관에서 개별적으로 이루어짐</li> <li>· 국토교통부 내에서도 주택정책과, 토지정책과, 녹색건축과, 주택건설공급과, 주거복지계획과, 주택기금과, 토지정책과, 도시정책과, 지적기획과 등 다양한 부서에서 관리됨</li> </ul> <div data-bbox="164 1009 763 1136"> </div> <ul style="list-style-type: none"> <li>· 부동산 보고통계 중 부동산 거래, 건축실적, 건축허가 및 착공 등의 통계는 RTMS, HIS, 세우터를 통해 시군구에서 직접 입력된 자료를 기반으로 됨. 기타 자료의 경우 엑셀 또는 전자문서 형태로 보고됨으로 입력과정과 문서작성과정에서 오류발생 가능성</li> </ul> <div data-bbox="164 1303 763 1479"> <div> RTMS HIS 세우터 </div> </div> <p>▶ <b>문제점 및 개선 필요</b></p> <ul style="list-style-type: none"> <li>· 개별적으로 통계를 생산함과 동시에 작성내용의 중복, 공표시기의 불일치, 통계 간 연계활용 부족 등의 문제</li> <li>· 보고통계 입력 오류로 인한 신뢰성 및 적시성 문제</li> <li>· 기술개발에 따른 부동산 정보시스템의 선진화 필요</li> </ul> | <p>▶ <b>보고통계의 입력오류 감소 및 적시성</b></p> <ul style="list-style-type: none"> <li>· 분야별 보고통계생산 시스템의 사용자 중심 환경 제공으로 업무효율화 및 통계생성시기 단축</li> <li>· 집계 및 입력과정에서 발생할 수 있는 오류가능을 줄이기 위한 "검증시스템 구축"기술</li> </ul> <p>▶ <b>통계연계 및 활용성 강화 및 중복 제거</b></p> <ul style="list-style-type: none"> <li>· 분산되어 있는 기관별 부동산 정보를 효율적으로 관리하기 위한 부동산 DW(data warehouse) 구축과 유관기관과의 연계를 위한 도구개발</li> <li>· 자료결합을 통한 코드표준화와 품질진단 체계마련으로 부동산 통계의 이용편의성과 효율성 증가</li> </ul> <p>▶ <b>빅데이터 기술 적용으로 효율성 및 무결성 보장</b></p> <ul style="list-style-type: none"> <li>· 부동산 자료의 수집 - 가공 - 적재를 위한 빅데이터 기술 적용으로 실시간 자료제공 및 시장분석 기반 마련</li> <li>· 무결성 자동검사 시스템 구축으로 자료의 오류감소</li> </ul> |

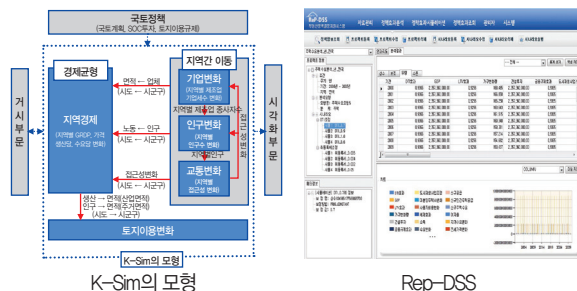


〈표 5〉 빅데이터 기반 부동산 시장 분석 및 예측 모형 개발

**현재(As-Is):** 부동산시장 환경 변화에 대응하기 위한 실시간 모니터링과 분석 및 예측 모형의 불확실성 관리와 빅데이터 기술 적용 미흡

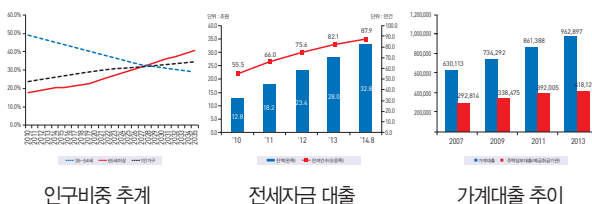
### ▶ 부동산 분석 및 예측 모형

- 부동산시장과 관련된 분석 및 예측모형은 부동산을 포함한 거시계량모형, 부동산 관련 정책효과 분석 모형, 헤도닉가격모형(Hedonic Price Model), 공간분석모형 등이 활용됨
- 부동산 정책에 활용되는 부동산시장분석 및 예측 모형은 EWS, K-REMAP, K-Sim, Rep-DSS 등이 운영



### ▶ 부동산시장 환경변화와 새로운 정책 자료 요구

- 저출산 고령화에 따른 부동산 수요의 변화와 맞춤형 부동산 정책에 대한 요구 증가
- 하우스푸어 및 렌트푸어 등 부동산 관련 사회문제 발생
- 가계부채의 증가에 따른 위험 관리의 요구와 부동산과 금융시장의 결합으로 인한 시장구조의 복잡화



### ▶ 문제점 및 개선 필요

- 부동산 정책지원을 위한 모형이 있지만, 모형간의 비교를 통해 모형의 최적화가 미흡
- 모형 간 통합 및 예측의 불확실성에 대한 관리체계 마련 필요
- 부동산시장을 둘러싼 환경의 급격한 변화에 대응하기 위한 부동산시장에 대한 실시간 진단체계 부재
- 비정형 빅데이터 기술의 부동산 분석 기술의 활용 방안 모색 필요

**미래(To-Be):** 빅데이터 기술기반 분석 및 예측모형 개발

### ▶ 부동산 분석 및 예측 모형의 최적화 및 불확실성의 관리체계

- 개별적 모형을 통해 분석 및 예측된 결과를 통합하고 최적화
- 추세 시뮬레이션을 수행하고, 이를 시각화함으로써 부동산시장의 수요, 공급, 가격 등과 관련한 고부가가치 정보를 탐색 및 산출
- 분석 및 예측된 결과에 대한 불확실성 관리와 감소를 위한 알고리즘

### ▶ 비정형 부동산 빅데이터 마이닝 체계 구축

- 부동산 관련 소셜 미디어, 뉴스 등 비정형 데이터의 수집 및 패턴의 실시간 모니터링과 분석 자료의 부족으로 정책에 반영

### ▶ 환경 변화에 선제적 대응을 위한 실시간 부동산시장 진단 및 모니터 체계

- 빅데이터를 접목하여 유용한 정보를 추출 가공하여 부동산시장을 실시간 모니터링 함으로 선제적 대응체계 마련
- 하부시장별로 국지적 불안을 조기 예측 및 이례현상 감지
- 부동산시장 안정화와 정책에 대한 효과를 주기적으로 모니터링 할 수 있는 능동적 대응체계
- GIS기술 발전에 따른 공간분석 기법과 부동산 빅데이터 분석 기술의 적용

### ▶ 부동산 빅데이터를 처리·분석·예측하기 위한 기반 여건 마련

- 부동산 빅데이터를 이용한 의사결정시스템(OLAP) 및 시뮬레이션 기술, 대규모분석을 위한 Scalability 확보기술, 방대한 양의 데이터에 대하여 고효율적인 행렬연산 시스템 등 기반 마련

〈표 6〉 빅데이터 기반 부동산시장 정부서비스 플랫폼 구축

| 현재(As-Is): 부동산 정보의 공개 및 지속가능한 대응시스템 마련을 위한 중·장기적 계획의 부재                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 미래(To-Be): 빅데이터 기술 기반 부동산시장 정부서비스 플랫폼 구축                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>▶ <b>부동산 정보 서비스</b></p> <ul style="list-style-type: none"> <li>· 현재 부동산 관련 정보는 공공부문의 경우 온나라부동산정보, R-one부동산통계정보시스템, APT2you 기타 통계청과 지방자치단체에서 서비스되고 있음. 민간부문은 부동산 114, 국민은행 등 있음</li> </ul> <div data-bbox="168 611 756 764">  <div> 온나라 R-one APT2you 서울시부동산정보광장 </div> </div> <p>▶ <b>부동산 통계협의회</b></p> <ul style="list-style-type: none"> <li>· 2013년 8월부터 각계각층에서 제기한 부동산통계의 생산 및 관리의 문제를 개선하고자 '부동산통계협의회'를 구성·운영 중임</li> </ul> <div data-bbox="171 950 756 1152">  </div> <p>▶ <b>공공데이터 공개 및 빅데이터 기술</b></p> <ul style="list-style-type: none"> <li>· 최근 정부는 국민 삶의 질을 높이고 가치를 향상하기 위한 창의적 아이디어와 기술의 융합·발전에 기초한 경제발전과 고용창출의 가능성을 위한 공공정보를 공개하고 있음</li> <li>· 과학기술을 통한 창조경제 기반 조성 빅데이터·초고성능컴퓨팅 등을 활용한 공공·민간 서비스 확대와 이를 통한 서비스산업 육성</li> </ul> <p>▶ <b>문제점 및 개선 필요</b></p> <ul style="list-style-type: none"> <li>· 개별적인 가격이 형성되는 부동산시장에 충분한 정보제공이 이루어지지 못해 합리적인 의사결정에 한계</li> <li>· 소비자의 이용성을 높이기 위한 편의성이 낮음</li> <li>· 관련 산업의 발전과 창조경제 실현에 부합하는 오픈 API 수준의 부동산 데이터가 공개되지 못함</li> <li>· 중·장기적인 측면에서 민간과 공공이 함께 참여하는 협력적 GOV 차원에서 부동산 정보서비스의 제공의 마스터플랜이 요구됨</li> </ul> | <p>▶ <b>사용자의 이용편리성 제고와 공공데이터를 위한 기반구축</b></p> <ul style="list-style-type: none"> <li>· 부동산 관련 분석 및 예측기능들을 정보 사용자 이용편이성에 맞게 서비스하기 위한 개발 및 제공</li> <li>· Open API기반의 개방형 정보서비스</li> <li>· 고급사용자를 위하여 WYSIWYG 방식으로 분석 및 예측을 수행 수 있는 모듈 제공으로 관련 분야의 정보수집 비용 감소</li> </ul> <p>▶ <b>부동산시장 분석을 위한 오픈소스 SW 제공</b></p> <ul style="list-style-type: none"> <li>· 대규모 분산 환경에서의 통계적인 분석을 위한 오픈소스 SW 기술 개발 및 부동산시장 관련 데이터에 특화된 SW 인프라 구축 및 개방으로 사회적으로 부동산시장 분석 비용 경감</li> </ul> <p>▶ <b>컨트론타워 운영방안 수립</b></p> <ul style="list-style-type: none"> <li>· 부동산시장 정보시스템 통합 및 유지 보수 계획 및 소프트웨어 품질보증 시스템 마련</li> <li>· 부동산시장 분석 및 예측 결과의 기관용, 대민용 부동산 정보 공개범위 및 산업화 전략 모색</li> <li>· 민간부문과 공공부문이 동시에 참여를 통한 지속가능한 부동산시장 대응체계 마련과 중·장기 발전 계획 제공</li> </ul> |

## V. 결론

빅데이터는 우리의 미래의 삶에 크고 작은 영향을 줄 것이다. IT 기술의 개발에 따른 지능형 과학의 발전과 생활의 편리성과 새로운 경제발전의 혁신적인 요소가 될 수 있을 것이다. 정부는 이러한 과학기술의 청사진에 기초하여 빅데이터 관련 분야에 대하여 육성·발전시키고 있다. 공공의 데이터를 개방하고, 빅데이터의 관련 성공사례를 발굴하는 등 사회문제 해결을 위한 다양한 형태의 시도를 하고 있다.

지금까지 부동산 정책은 동일한 시장을 놓고 경기활성화와 가격 안정이라는 두 가지 상충된 가치 사이에서 정책수단들이 선택되면서 정책의 일관성과 신뢰성이 저하되었고 과잉 및 과소대응의 문제와 함께 사회·경제적 부작용이 발생하였음은 주지의 사실이다. 따라서 경기 및 금융시장 활성화와 주거복지, 가격안정이라는 다양한 가치를 조화시킬 수 있는 방법이 모색되어야 한다. 이를 위해 부동산 산업과 국민경제간의 연관관계와 시너지파악을 위한 기술개발이 필수적이라고 판단된다.

부동산시장의 선진화 및 안정화를 기하기 위하여 우선 지방자치단체 차원에서 부동산 정보 수집 및 생성 체계에서 발생하는 오류가능성을 개선함으로써 원천자료의 신뢰성을 강화 해야만 한다. 이러한 원천자료의 신뢰성을 확보하지 못한다면 아무리 공고한 차후의 분석 및 모형 개발은 사상누각에 불과할 것이 자명하다. 다음 단계에서 이를 효율적 지원 관리 할 수 있는 운영체계를 마련함으로써 통계생산체계를 고도화 하고, 부동산지표에 대한 국민 체감형 부동산 통계의 생성이 요구되어 진다. 부동산시장 분석 및 예측에 필요한 통계지표를 평가하고, 부동산시장 변화에 대응하기 위한 부동산 통계수요의 발굴 및 구축, 공간정보 분야의 통계자료의 부동산시장 분석 및 예측의 활용성을 증대시켜야 한다. 소셜 미디어, 뉴스, 뉴스 댓글 등의 비정형 데이터는 통계 수치보다 더욱 빠른 반응을 보이기 때문에 부동산 정책결정을 위한 실시간 지원이 가능토록 하여야 한다. 따라서 비정형데이터 수집 및 처리를 통

한 부동산시장 모니터링 기반 마련 또한 간과되어서는 안 될 매우 중요한 부문 중에 하나이다.

최적화된 부동산시장 분석 및 예측 모형 개발에 있어 부동산시장 조기경보시스템(EWS; Early Warning System), 부동산시장 진단 및 전망 시스템(K-REMAP; The KRIHS System for Real Estate Market Analysis & Preestimate), 부동산정책결정지원시스템(Real Estate Policy-Decision Support System; Rep-DSS) 등에서 시계열모형, 계량경제모형, 시스템 다이내믹스 및 GIS기법 등이 개별적으로 활용되고 있다. 하지만 예측모형별로 오차발생에 대한 불확실성과 공간적 상호효과에 대하여 간과하고 있는 실정이다. 따라서 기존 부동산 예측모형의 고도화와 더불어 새로운 부동산 환경에 대응하기 위한 공간적 특성을 고려한 예측모형의 개발이 요구되어 진다.

빅데이터 기술의 정부 정책과정에 도입은 과학적 정책결정 및 집행을 통해 행정적 행정과 주민편의를 높일 수 있을 것이다. 하지만, 분석결과에 사로잡혀 오히려 주민과의 대면접촉이 감소되어 행정의 대응성을 감소시키는 결과를 낳을 수도 있다. 신뢰성이 낮은 빅데이터 수집 및 분석은 정책결정의 오류 또는 노이즈로 작용할 가능성이 높다. 이러한 오류를 감소시키고 빅데이터 기술의 행정과정에 성공적으로 도입 활용되기 위해서는 무엇보다 빅데이터 활용과 관련된 전문 인력 양성이 절대적으로 요구되고 있다. 향후에는 얼마나 많은 빅데이터 전문 인력을 확보하고 있느냐가 핵심적 경쟁력이 될 것이다. ☺

### 〈참고문헌〉

- 국토교통부(2014.1.23.), “정부, 올해 385개 공간정보사업 추진, 2,946억 원 규모” 국토교통부 보도자료  
김대중(2014), 정부3.0을 위한 공간빅데이터 구축 및 활용방안, 『월간국토』, 387: 42-51  
김동원(2013), 빅데이터의 분야별 활용사례, 『경영논총』,

- 34: 39-52
- 김민수(2014), 국내외 공간빅데이터 정책 및 기술동향, 「월간국토」, 389: 30-39
- 김상락·강만모(2014), 빅데이터 분석기술의 오늘과 미래, 「정보과학회지」, 32(1): 8-17
- 김희수(2014), 공간빅데이터 체계구축·활용 정책방향, 「월간국토」, 389: 6-11
- 서수복(2006), 가격예측모형에 의한 부동산 정책평가, 진주대학교 박사학위 논문
- 안창원·황승규(2012), 빅데이터 기술과 주요이슈, 「정보과학회지」, 30(6): 10-17
- 오라일리미디어(2013), 「오라일리 심층보고서 빅데이터 어떻게 활용할 것인가?」, 제이펍
- 이영주 외(2014), 복지·안전국토 실현과 공간빅데이터 활용, 「월간국토」, 389: 12-21
- 이정미(2013), 빅데이터의 이해와 도서관 정보시스템의 활용, 「한국비블리아학회지」, 24(4): 53-73
- 최진원·김이연(2012), 빅데이터시대에 효과적인 시각커뮤니케이션을 위한 인포그래픽 연구: 정부부처 및 공공기관을 중심으로, 「한국과학예술포럼」, 11(1): 165-175
- 한국정보화진흥원(2012), 「빅데이터로 진화하는 세상 : Big Data 글로벌 선진사례」, 한국정보화진흥원
- 한국정보화진흥원(2013), 「새로운 미래를 여는 빅데이터 시대」, 한국정보화진흥원
- Gantz, J. and Reinsel D(2011), Extracting value from chaos, IDC I-View, June 2011, <http://www.emc.com>
- Gartner(2011), Pattern-Based Strategy : Getting Value from Big Data, Gartner Special Report(Jun 2011)
- IDC(2012), Worldwide Big Data Technology & Services 2012-2015 Forecast, December 2012, IDC #238746, Volume: 1
- Kobayashi(2013), Real estate market & housing policy in Japan
- Manyika et al(2011), Big Data: The Next Frontier for Innovation, Competition, and Productivity, Technical Report, McKinsey Global Institute
- Newman, D(2011), How to Plan, Participate and Prosper in the Data Economy, Gartner, Stamford, CT