https://doi.org/10.15813/kmr.2019.20.3.005

서울시 공공빅데이터 활성화 방안 연구[1]

A Study on Policies to Revitalize the Public Big Data in Seoul

최 봉 (Bong Choi) 서울연구원 시민경제연구실[2]
윤 종 진 (Jongjin Yun) 서울연구원 시민경제연구실[3]
엄 태 휘 (Taehyee Um) 경희대학교 호텔관광대학[4]

ABSTRACT

The purpose of this study is to investigate the current state of public Big Data in Seoul and suggest policy directions for the revitalization of Seoul's public Big Data. Big Data is perceived as innovation resources under the era of 4th Industrial revolution and Data economy. Especially, public Big Data serves a significant role in terms of universal access for citizens, startup, and enterprise compared with the private sector. Seoul reorganized a substructure of government's focus on Big Data and established organizations such as Big Data Campus and Urban Data Science Lab. Although the number of public open Data has increased in Seoul, there exists not much Data with characteristics similar to Big Data, such as volume, velocity, and value. In order to present the direction of Big Data policy in Seoul, we investigate the current status of Big Data Campus and Urban Data Science Lab operated by Seoul City. Considering the results of this study, we have proposed several directions that Seoul can use in establishing big data related strategies.

Keywords: Public bigdata, Seoul open data portal services, Big data campus, Policy direction, Seoul

---

1) 본 연구는 제1저자와 제2저자의"서울시 공공빅데이터 활성화방안 (서울연구원, 2018)"의 연구보고서를 재구성하였음.
논문접수일: 2019년 8월 5일 1차 수정: 2019년 9월 5일; 게재확정일: 2019년 9월 16일
2) 제 1저자 (cbong@si.re.kr)
3) 제 2저자(jjyun@si.re.kr)
4) 교신저자 (thum@khu.ac.kr)

1. 서론

초연결성(hyper-connected), 초지능화(hyper-intelligent)라는 키워드로 대변되는 4차 산업혁명은 정보통신기술(ICT)을 통해 인간과 사물이 상호 연결되고 보다 지능화된 사회로의 변화를 가져온다. 이에 근간을 이루는 핵심 기술 중 하나가 빅데이터(bigdata)이다. 빅데이터는 100만 기가바이트와 동일한 크기인 페타바이트(petabyte) 이상의 크기를 가지고 새로운 가치 추출을 위해, 기존의 기술 또는 기법으로 처리하기 어려운 규모(volume), 다양성(variety), 신속성(velocity), 가치(value) 그리고 진정성(veracity)과 같은 특징을 갖는 데이터를 지칭한다(Baesens et al. 2016).

미국을 비롯한 여러 국가들은 민간뿐만 아니라 공공부문에서도 빅데이터의 다양한 활용방안을 마련하고 있다(Chen and Zhang 2014). 대한민국 중앙정부 역시 과학기술정보통신부, 국토교통부, 통계청 및 국세청 등에서 빅데이터를 활용한 서비스 및 사업화 전략을 모색하고 있다(정보통신정책연구 2017). 이러한 흐름 속 서울시는 빅데이터를 4차 산업혁명시대의 새로운 디지털 자원으로 인식하면서 이와 관련한 공공데이터 정책을 수립해 나가고 있다. 서울시는 2015년 이후 공공데이터 제공 및 이용활성화 시행계획을 추진 중이며, 2018년도에 처음으로 서울시 빅데이터 실행계획을 수립하여 데이터 기반 사회혁신 생태계 조성에 힘쓰고 있다(서울특별시 2018). 또한 정책의 주요 수단으로 통계와 원천데이터를 제공하는 서울열린데이터광장과 서울빅데이터 캠퍼스를 운영함으로써 서울시에서 공급하는 공공데이터의 양적 확대를 추구하고 있다.

그러나 서울시 공공데이터의 양적 확대에도 불구하고 민간부문의 공공빅데이터 활용 및 사업화는 아직까지 저조한 실정이다. 구체적으로 공공데이터법 제정 이후 서울시의 공공데이터 누적 개방건수는 약 7.2배 증가(2013년에 684건에서 2018년 4,892건으로 증가)하였으나 2017년 민간부문의 활용성과는 24건에 불과하며, 이마저도 교통분야 18건, 대기환경 4건 등으로 일부 분야에만 집중되어 있다(서울특별시 2018). 뿐만 아니라 데이터의 생애주기(생성·수집·저장관리·분석·활용) 단계별로 민간부문이 참여할 수 있는 비즈니스 기회가 있음에도 불구하고 현재는 분석 및 활용단계에 초점이 맞춰져 있어, 민간 참여형 빅데이터 생태계 구축 및 육성에서 어려움을 가지고 있다(성욱준 2016).

기존 빅데이터의 정책과 관련한 연구를 살펴보면 공공영역에서 활용가능한 빅데이터 적용방안을 경기도를 대상으로 연구하였고(유영성 등 2014), 권영일 외(2017)은 글로벌 빅데이터 정책 현황을 통해 국내 빅데이터 정책 수립 방향을 개괄적으로 제시하였다. 배동민 외 (2013)의 연구에서는 국내와 국외 빅데이터 관련 시장 실태 및 정책 동향을 파악함으로써 시장 활성화를 위한 실무적 제언을 제시했다. 또한 재난관리 분야, 환경 분야 및 공공부문 등 빅데이터가 활용될 수 있는 다양한 분야에 초점을 맞춘 정책 연구도 이루어졌다(김도우 2017; 신동희·김용문 2015; 임은선 등 2018). 한편 서울시 빅데이터 관련 연구는 서울시에서 발생하여 적립된 빅데이터를 기반으로 분석을 하여 사회적 현상에 대한 고찰을 한 연구들이 있었다(강진동 2016; 김민길 등 2018). 그러나 이러한 연구들은 포괄적이고 일반적 형태의 빅데이터 발전방향을 제시하거나 특정 산업군에 초점을 맞춘 시사점을 제시하였다. 또한 정책적 제언 역시 새롭게 시행되는 정책들에 대한 충분한 고려가 부족하다는 한계점을 가진다. 더불어, 서울시라는 특정 지역에 초점을 맞춘 관련 연구가 없었다. 때문에 서울시의 공공빅데이터 현황 분석과 이를 토대로 정책적 제언을 하는 것은 앞으로의 빅데이터 관련 정책수립 및 활성화 방안에 대한 중요한 자료로 활용될 수 있을 것이다.

따라서 본 연구의 목적은 서울시의 공공빅데이터 실태를 파악하여 서울시 공공빅데이터 활성화를 위한 정

책 방향을 제언하는 것이다. 연구의 구성은 다음과 같다. 첫째, 문헌연구를 통해 공공빅데이터의 특징과 역할을 알아보고 서울시의 빅데이터 관련 정책을 살펴본다. 둘째, 서울시가 운영하는 열린데이터광장과 빅데이터캠퍼스의 공공빅데이터 실태를 통해 데이터 생애주기 단계별 데이터 제공과 활용의 촉진 및 저해요인을 파악한다. 셋째, 선행연구와 실태조사를 기반으로 제공자 및 이용자 측면의 활성화 방안을 모색하고, 서울시의 공공빅데이터 활성화를 위한 기본 방향을 제시한다.

2. 기존문헌 연구

2.1 빅데이터와 공공빅데이터

빅데이터는 정보통신기술 발전과 스마트기기의 확산, 사물인터넷(Internet of things)의 확대 및 SNS(social network service) 활성화에 따른 데이터의 증가로 주목받기 시작하였다. 2012년 전후로 맥킨지, 가트너 그리고 삼성경제연구소 등의 여러 연구기관에서 빅데이터와 관련한 연구를 시행하면서 빅데이터는 공공 및 기업에서 정보화 시대의 새로운 혁신 자원으로 인식되기 시작했다. 국내 데이터산업의 시장 규모 역시 2010년 8조원에서 2017년 14조원에 육박하여 연평균 7.5%의 성장세를 유지하면서 관련 비즈니스 분야에 지대한 영향을 미치고 있다(한국데이터진흥원 2018).

빅데이터의 선행연구에서 관련 정의를 살펴보면 문헌마다 조금씩 차이를 가지지만 공통적으로 '대용량 데이터'를 일컫는다. 최근에는 데이터 자체를 말하는 동시에 데이터의 처리, 축적, 분석 기술 혹은 조직 등 빅데이터를 둘러싼 생태계 내 구성요소까지 포함하는 개념으로 확장되고 있다. 빅데이터의 기존 연구가 밝힌 정의는 다음 <표 1>과 같다.

빅데이터는 규모, 다양성 그리고 신속성의 대표적인 특징 세가지를 가지며(Laney 2001), 이후 빅데이터의 활용분야 혹은 방식 등에 따라 가변성(variability), 정확성, 가치 또는 시각화(visualization) 특징을 가지기도 한다. 한편, 빅데이터는 데이터의 생성과정과 데이터 활용 목적에 따라 정형, 반정형, 비정형 데이터로 구분할 수 있다. 정형(structured)데이터는 표와 같이 고정된 열에 저장되는 데이터와 파일을 의미하며 지정된 행과 열에 의해 속성이 구별되는 형태이다. 반정형(semi-Structured)데이터는 데이터의 구조에 대한 메타정보를 갖고 있는 형태를 말하며, OpenApi, 사물인터넷 센서데이터 등이 여기에 속한다. 비정형(unstructured)데

<표 1> 빅데이터 정의

기관	정의
Mckinsey(2011)	일반적인 데이터베이스 SW가 수집·저장·관리·분석할 수 있는 처리능력을 초과하는 규모의 데이터집합
IDC(2011)	다양한 형태의 방대한 데이터로부터 고속 캡처, 데이터 탐색 및 분석을 통해 경제적으로 필요한 가치를 추출하기 위해 설계된 차세대 기술 및 아키텍처
Gartner(2012)	더 나은 의사결정, 시사점 발견 및 프로세스 최적화를 위해 사용되는 새로운 형태의 정보처리가 필요한 대용량, 초고속 및 다양성의 특성을 가진 정보자산
교육과학기술부 등(2012)	데이터 형식이 다양하고 생성 속도가 매우 빨라, 새로운 관리 및 분석 방법이 필요한 대용량 데이터
정보통신기술협회(2015)	새로운 가치 추출을 위해 기존의 기술 또는 기법으로 처리하기 어려운 특징(규모, 신속성, 가변성 다양성, 진정성)을 갖는 데이터 모음

출처 : 토지주택연구원(2015)의 연구를 재구성

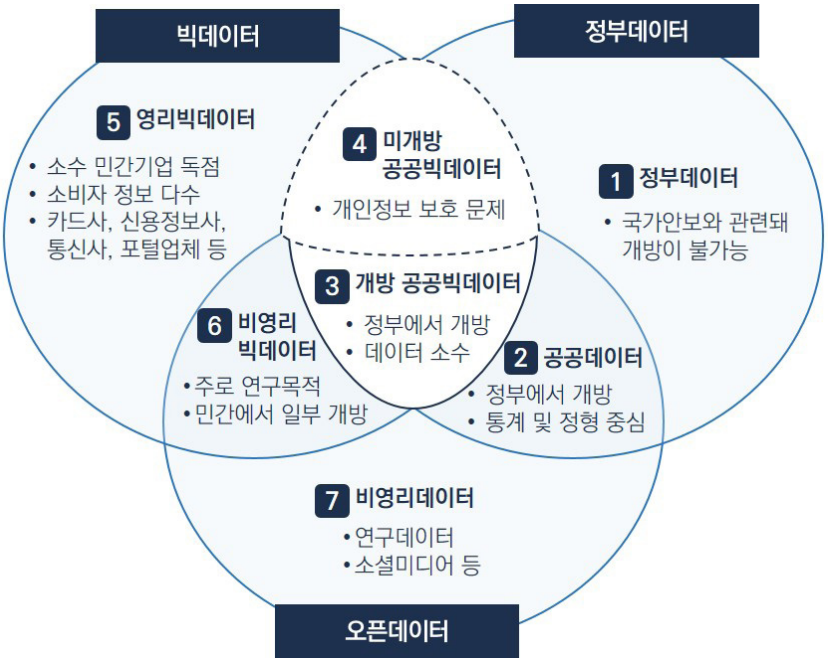
이터는 텍스트, 이미지 혹은 동영상과 같이 사전에 정의된 구조가 없는 형태를 말한다.

최근에는 빅데이터가 기계학습 및 딥러닝 등 인공지능의 알고리즘 구현을 위한 학습데이터로 이용되면서, 활용성이 낮았던 소리·이미지·비디오 등 비정형데이터에까지 분석대상이 확장되고 있다(Ko et al. 2014). Mckinsey(2018) 보고서에 따르면 빅데이터 분석 방식은 분석의 대상인 문제의 유형에 따라 분류(classification), 지속적인 추정(continuous estimation), 군집화(clustering), 최적화(optimization), 이상치 탐지(anomaly detection), 순위 매기기(ranking), 추천(recommender systems) 그리고 데이터 생성(data generation)으로 구분한다. 맥킨지가 약 400개 기업의 사례를 통해 빅데이터 분석 기술 적용분야를 조사한 결과, 분류(72%)가 가장 많이 적용될 수 있으며 지속적인 예측(66%), 군집화(55%), 최적화(37%), 이상치 탐지(24%), 순위(17%), 추천(15%) 마지막으로 데이터 생성(7%) 순으로 나타났다.

빅데이터의 개념적 모호성으로 인해 공공빅데이터는 실체를 가진 개념이라기 보다 오픈데이터, 정부데이터, 빅데이터의 성격을 가진 '공공부문에서 공개한 빅데이터'에 가깝다.

먼저, 오픈데이터는 목적에 관계없이 누구나 어디에서 자유롭게 접근이 가능하고 공유 및 활용할 수 있는 데이터를 의미한다. 저작권과 같은 제한적인 규칙을 적용하지 않기 때문에 법적 개방성의 특징을 가지며, 특정 소프트웨어에서만 확인이 가능한 형태가 아니기 때문에 기술적 개방성을 갖는다(김구 2017).

공공데이터는 정부데이터 가운데 오픈데이터 성격을 갖는 데이터로 "데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 전자적 방식으로 처리한 자료 또는 정보"이다(공공데이터전략위원회 2016). 2013년 '정부3.0'을 추진하면서 제정된 공공데이터법에 따



<그림 1> 빅데이터, 오픈데이터 및 정부데이터에 따른 유형

라 국가 및 지방자치단체는 '공공데이터의 제공 및 이용 활성화에 관한 기본계획'을 수립하고 관련 정책을 추진하고 있다. 정부의 데이터 개방이 원활하게 이루어질 경우 새로운 제품과 서비스에서 데이터의 재활용이 이루어질 수 있다. 사회적 가치 실현에 있어 혁신적인 솔루션을 탐색할 수 있고, 공공행정 부문에 있어 데이터를 공유하여 운영의 효율성을 증대시킬 수 있다. 더불어 사회적 데이터 활용과 관련한 정치 및 사회생활에 시민참여를 촉진시킬 수 있고 나아가 정부의 투명성을 제고할 수 있다(관계부처합동 2018).

데이터는 생성주체와 공개여부에 따라 구분하여 다음 <그림 1>과 같이 유형화할 수 있다(Joel Gurin 2014). 이를 종합해본다면 공공빅데이터는 영리빅데이터 수준의 가치를 가지고 있을 것으로 보인다. 다만 개인정보 등과 같은 제약조건으로 인해 개방되지 못하거나, 개방되더라도 소수에 그쳐 아직은 활용도가 떨어질 수 밖에 없다. 그럼에도 미개방 공공빅데이터를 개방하고 영리빅데이터의 거래 활성화를 지속적으로 유도한다면 접근성 및 개방성이 개선될 수 있을 것이다.

공공빅데이터는 다양한 이해당사자가 참여하여 활용하기 때문에 민간 대비 관련 산업 생태계에 미치는 파급효과가 크다. 민간 영리빅데이터는 통신사 카드사 신용정보사 포털업체 등 소수의 민간기업이 정보를 독점하고 있어, 활용성이 높음에도 접근과 개방성이 취약한 상황이다. 앞으로도 기업들이 데이터를 그들만의 자산으로 인식하여 개방에 있어 소극적인 자세를 취할 것이다. 이러한 환경 속에서 공공빅데이터는 빅데이터 획득에 대한 차별을 제거해줄 수 있고 관련 산업 발전의 거대 인프라 역할을 수행할 수 있다.

2.2 공공빅데이터 관련 선행연구

빅데이터 연구는 2011년부터 본격적으로 진행되었다. 초기 빅데이터에 대한 연구는 외국의 빅데이터 활용사례 연구, 빅데이터 분석 및 활용방안 등이 대부분을 차지했다. 빅데이터 등장에 따른 활용방안과 대응전략 등에 대한 논의가 이루어졌고(김성태 2011), 기업 및 국가 유관기관의 빅데이터 관련 보고서 등이 등장했다(조성우 2011; 이각범 2011). 이후 빅데이터의 특징, 빅데이터 관련 기술에 대한 소개 및 응용사례 연구가 이루어졌다(Kim 2012). 특히 빅데이터의 선도적 위치에서 활용을 하고 있는 나라의 사례 연구와 함께 국내 빅데이터 활성화 방안을 논의하는 연구가 많이 나타났다(임용재 등 2012; 조영임 2013). 빅데이터에 대한 연구가 활성화 되면서 구체적인 영역 안에서의 빅데이터 활용방안 혹은 전략적 접근에 대한 연구들이 이루어졌다(김동완 2013). 한편, 빅데이터에 대한 연구가 시작되면서 개인정보 및 데이터 보안에 대한 논의 역시 지속적으로 이루어졌다(이창범 2013; 박민정·채상미 2017).

공공빅데이터 연구는 공공에서의 빅데이터 구축의 중요성과 빅데이터 구축을 위한 방안(이강용 등 2012), 국내외 공공부문에서의 빅데이터 활용 실태 분석(정지형 김강훈 2012), 빅데이터를 통한 공공 서비스 설계(표순희 등 2015), 공공빅데이터 정책 사례 연구 및 제언에 대한 연구가 지속적으로 이루어졌다(성욱준 2016; 오재인 2017; 김도우 2017). 선행 연구에 따르면 빅데이터의 공공부문 활용도는 광범위하며 여러 분야에 걸쳐 서비스 활성화에 기여할 수 있는 잠재력을 가진다고 했다(김동성 등 2017). 또한 공공 빅데이터를 통해 대민 서비스 개선, 사업체에서는 빅데이터를 활용한 사업활성화 방안을 모색할 수 있을 뿐만 아니라 공공분야 운영의 효율성 및 투명성을 제고할 수 있을 것이라고 강조했다. 이러한 연구들을 종합해 살펴봤을 때, 공공 빅데이터에 대한 많은 연구가 진행되었으나 서울시가 주관하는 특정 공공 빅데이터 활성화 운영실태를 파악하여 해당 지역에 초점을 맞춘 정책 제언을 제시한 연구는 없었다. 따라서 본 연구는 서울시의 공공 빅데이터 정책을 살펴보고, 운영 실태를 파악하여 정책 수립에 대한 방향성을 제시하고자 했다.

2.3 서울시의 빅데이터 정책

공공데이터 개방은 2009년 서울·경기지역의 버스정보 데이터를 이용해 버스 노선과 도착을 알려주는 '서울버스' 어플리케이션을 개발하여 서비스를 하면서 공공데이터의 이용과 권리, 활용에 대하여 논의되기 시작하였다. 이후 중앙정부는 2013년에 '정부 3.0'을 발표하면서 공공데이터법을 제정하여 본격적인 정책을 시행하였다(관계부처합동 2013). 서울시의 빅데이터 정책은 2013~2015년 빅데이터의 저변 확대를 위한 탐색과 조직개편 시기, 2016~2017년 빅데이터 활성화를 위한 개별 사업이 추진된 시기, 2018년 제도적 기반 마련과 서울시 빅데이터 실행계획이 수립된 시기로 구분할 수 있다.

2013년부터 데이터의 중요성을 본격적으로 논의하면서 관련 정책 추진을 위한 탐색적 사업이 추진되었고, 데이터를 통한 서울시의 문제해결 시범과제를 제안했다. 이 시기 KT 유동인구데이터와 택시 승·하차 데이터를 활용한 통행량을 분석해 노선수요가 높은 지역을 중심으로 심야버스 노선을 제안한 '올빼미 버스'가 시행되었다. 2015년에는 내부직원용 빅데이터 분석 인프라를 구축하여 데이터 기반 행정의 확산을 추진하였다. 2016년에는 서울특별시 정보화전략위원회에서 5개년 기본계획인 '서울디지털기본계획2020'을 수립하였으며(서울특별시 2015), 빅데이터 분야에서는 '데이터기반 사회혁신 기반 조성'과 '빅데이터 공유·활용 플랫폼 고도화' 등의 전략을 추진하였다. 또한 빅데이터캠퍼스를 통해 데이터 확보가 어려운 시민·대학·중소상공인 등의 디지털격차를 해소하는 목표를 설정했다. 그리고 정책효과 분석과 입지 및 위치 분석 등을 활용한 각종 빅데이터 관련 개별 사업이 추진되었다. 2018년에는 '서울시 빅데이터 시행계획'을 수립하여 빅데이터 캠퍼스 운영, 과학행정 정착 및 내재화, 민관 분석과제 발굴, 빅데이터 교육사업을 주요 사업으로 구성하였다.

이처럼 중앙정부의 공공데이터 관련 정책 추진과 함께 서울시 공공데이터 활성화 방안이 단계적으로 확장됨에 따라 공공데이터의 개방 역시 증가하고 있다. 서울시의 공공데이터의 개방 건수는 684건(2013년)에서 4,892건(2018년 8월)으로 약 7.2배 증가하였고, 빅데이터 활용을 지원하기 위해 열린데이터광장 및 빅데이터 캠퍼스를 설치하여 운영 중에 있다. 또한 서울시는 서울대학교와 공동으로 도시데이터사이언스 연구소를 개소하여 분석, R&D, 교육 등 다양한 사업을 통해 빅데이터의 인식을 확산하고 있다. 그럼에도, 활성화의 밑바탕이 데이터 그 자체이기 때문에 서울시 빅데이터 정책의 주된 방향은 공공데이터의 양적 개방이라 볼 수 있다.

<표 2> 서울시 주요 빅데이터 정책

년도	내용
09	• 서울·경기지역의 버스정보 데이터를 이용해 버스 노선과 도착을 알려주는 '서울버스' 어플리케이션을 개발하여 서비스를 하면서 공공데이터의 이용과 권리, 활용에 대하여 논의
13	• '정부 3.0'을 발표를 통해 공공데이터법 제정 • KT 유동인구데이터와 택시 승·하차 데이터를 결합, 통행량을 분석하여 심야버스 노선을 제안한 '올빼미 버스' 시행
15	• 내부직원용 빅데이터 분석 인프라를 구축하여 데이터 기반 행정처리 환경 조성
16	• 서울특별시 정보화전략위원회 주관 5개년 기본계획인 '서울디지털기본계획2020' 수립 • '데이터기반 사회혁신 기반 조성'과 '빅데이터 공유·활용 플랫폼 고도화' 등의 활성화 전략 추진 • 데이터 확보가 어려운 대상군을 위해 빅데이터캠퍼스를 설립하여 디지털격차를 해소하고 함
18	• 빅데이터 캠퍼스 운영, 과학행정 정착 및 내재화, 민관 분석과제 발굴, 빅데이터 교육사업을 주요 사업 내용으로 구성한 '서울시 빅데이터 시행계획' 수립

3. 연구방법

3.1 분석 절차

서울시 열린데이터광장에서 제공하는 공공데이터의 실태를 파악하기 위해 데이터의 생애주기 단계별로 구분하여 특성을 분석하여 규모·신속성·가치 등 빅데이터 특성에 대한 가능성을 확인하였다. 데이터 생애주기는 데이터의 생성-활용-폐기 또는 재활용까지의 일련의 과정을 의미하며, 본 연구에서는 생성-수집 및 저장관리-분석-활용으로 구분하였다. 데이터의 생애주기는 단계별로 다음과 같은 특징을 가진다(성욱준 2017). 우선 생성 단계에서는 사실 또는 현상을 관찰해 자료 또는 정보를 최초로 생성하는 과정이다. 데이터를 생성하는 주체로부터 정형, 반정형 혹은 비정형의 형태로 생성이 된다. 이와 같은 데이터는 사회네트워크, 공공·민간 사업정보 또는 사물인터넷으로부터 창출될 수 있다. 수집 및 저장관리 단계는 내부 및 외부에 존재하는 데이터를 일괄 혹은 실시간으로 수집하고 데이터를 안전하게 활용하기 위해 저장하고 관리하는 과정이다. 이 단계에서 데이터는 파일 및 데이터베이스의 형태로 저장이 된다. 분석 단계는 원시 데이터 또는 분석결과를 서비스로 제공 및 이용하는 단계로서, 분석된 데이터는 파일, 시트, 차트, 링크 혹은 OpenApi 등으로 전달될 수 있다. 활용 단계는 분석결과를 의사결정 프로세스에 반영해 경제적 혹은 사회적 효용을 높이는 과정이다. 데이터를 활용함에 있어 빅데이터에 대한 규모 측정은 데이터의 관측치 수로 확인을 한다. 신속성은 데이터의 갱신주기로 확인을 할 수 있고, 가치는 데이터의 다운로드 및 호출건수로 측정해볼 수 있다.

3.2 연구 대상 및 분석 단위

분석의 대상은 2018년 8월 기준 '서울시 열린데이터광장 서비스현황' 메타데이터다. 메타데이터(Metadata)는 '데이터에 대한 데이터'로 데이터에 관한 정보의 기술, 데이터 구성의 정의, 데이터 분류 등 자료의 속성을 설명하기 위한 정보를 의미한다(정보통신용어사전). 빅데이터캠퍼스의 현황 파악을 위해서는 서울시에서 운영중인 상암 빅데이터캠퍼스의 프로그램 현황, 이용자의 특징과 데이터 활용 방안 등을 조사하였다. 빅데이터캠퍼스의 운영 프로그램은 연차별 사업계획서와 홈페이지를 확인하였다. 분석 대상 프로그램은 <표 3>에 나타나있다.

<표 3> 빅데이터캠퍼스 협치 모델 및 운영 서비스

교육 프로그램 협치	
대학교 정규 강의 연계	파트너쉽 대학교의 빅데이터 강의 및 실습 정규과정 운영
오픈 강의	파트너십 기업, 연구소의 일반인 대상 빅데이터 강의 및 실습 과정 운영
기업 교육 연계	파트너십 기업 빅데이터 교육과정 운영
파트너십 협치	
데이터 협력	보유하고 있는 데이터를 캠퍼스에서 활용 가능하도록 제공
빅데이터 분석 협력	서울시민에게 도움이 되는 분석과제 공동 분석 프로젝트 진행
빅데이터 공모전	서울시와 협력기관간 공모전을 공동으로 주최하거나 상호 후원
교육 프로그램 진행 및 일자리 창출	빅데이터 전문인력 양성을 위해 캠퍼스에서 협력기관의 교육 프 로그램 진행 또는 운영 지원

4. 분석 및 결과

4.1 서울시 열린데이터광장 현황과 이용실태

데이터의 생성 단계를 살펴보면 다음과 같다. 공공데이터는 2018년 8월까지 4,892건이 개방되었으며 보건(1,688), 일반행정(596), 문화관광(578) 순으로 높게 나타났다. 한편 도시관리(183), 교육(100), 안전(99) 순으로 적은 개방건수를 보였다. 2013년 대비 2018년의 전체 공공데이터의 누적개방건수를 비교했을 때 약 7.2배 증가한 것으로 나타났다. 분야별로는 교육(25배), 보건(15.6배), 복지(12.6배) 순으로 높은 증가를 보였다. 교통(4.3배), 안전(3.3배), 환경(2.2배) 분야는 증가율이 다소 낮은 것으로 확인되었다.

데이터 공급자는 서울시가 86%, 중앙정부 및 공공기관이 9% 등으로 대부분 서울시 및 산하기관·본청·사업소가 주를 이루는 것으로 나타났다. 한편, 데이터 제공자로서는 자치구가 48%로 가장 높았으며, 서울시 정보기획관(30%), 서울시 전 부서(13%), 서울시 공사·출연기관(4%) 등 순으로 나타났다. 공공데이터의 유형은 원천 데이터의 비율이 약 69%로 가장 높았으며 통계 데이터는 24%, 공간데이터는 7% 정도로 나타났다.

수집 및 저장관리 단계에서 공공데이터의 저장형태를 살펴보면 데이터베이스가 약 83%, 파일 형태가 약 17%로 나타났다. 그러나 데이터 공급자의 유형별로 살펴봤을 때 서울시 전 부서 및 본부와 사업소에서는 파일을 통한 수집 비율이 높게 나타나 아직은 시스템 간 연동이 미흡한 것으로 보여진다.

분석 단계에서 공공데이터의 서비스에 활용되는 추출물의 비중은 오픈Api로 서비스하는 비율이 약 92%로 가장 높은 활용도를 보였다. 구체적으로 오픈Api의 활용 영역에 대해 살펴보면 보건(97.4%), 문화관광(94.8%), 환경(93.2%), 산업경제(92.8%)는 오픈Api 서비스 비율이 90% 이상으로 높게 나타났고, 교통(68.9%) 분야는 가장 낮은 활용도를 보이는 것으로 확인되었다.

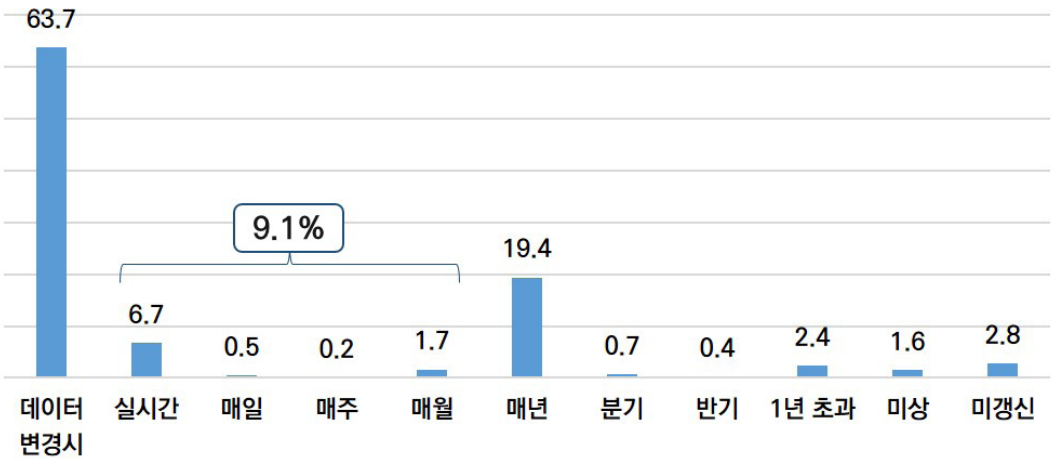
공공데이터의 활용단계를 살펴보면 다음과 같다. 먼저 규모의 측면을 살펴보면, 샘플링데이터 197개에 대하여 관측치 1,000건 미만의 공공데이터가 전체의 71.6%로 높은 비율을 보였으며, 약 9%만이 10,000건을 넘어서는 것으로 나타났다. 따라서 전체적으로 봤을 때 빅데이터 활용가능성은 낮은 것으로 볼 수 있다. 신속성의 측면을 살펴봤을 때, 공공데이터의 9.1%만이 한 달 이내에 정기적인 업데이트가 이루어졌고, 이중 6.7%만이 실시간 업데이트가 이루어지는 것으로 나타났다. 구체적인 공공데이터 갱신주기는 <그림 2>와 같다.

데이터 활용의 가치 측면을 살펴봤을 때, 공공데이터 다운로드 비율 중 교통 분야가 73.9%, 환경 분야가 25.3%로 두 분야가 99.2%를 차지하며 지배적인 분포를 보였다. 이밖에 문화관광, 산업경제, 안전 등 다른 분야는 상대적으로 활용도가 떨어지는 것으로 나타났다. 구체적으로 교통분야에서는 지하철 도착 및 위치, 도로소통, 버스 정류장·노선·위치, 나눔카예약가능정보 등 8개 영역에서 데이터 활용이 주로 이루어졌다. 환경 분야에서는 실시간 대기환경과 오존·황사·미세먼지 대기오

<표 4> 분야별 공공데이터 누적개방건수

구분	보건	문화관광	일반행정	산업경제	환경	복지	교통	도시관리	안전	교육	계
13년(A)	108	95	109	79	143	25	53	38	30	4	684
18년(B)	1,688	578	596	489	316	315	277	183	99	100	4,892
B/A	15.6	6.1	5.5	6.2	2.2	12.6	4.3	4.8	3.3	25.0	7.2

단위: 13년(A)와 18년(B)는 건, B/A는 %



단위: %

<그림 2> 공공데이터 갱신주기

<표 5> 분야별 공공데이터 다운로드 수, 비율, 데이터 1개당 평균 다운로드 수

구분	다운로드 수	비율	데이터 1개당 다운로드 수	구분	다운로드 수	비율	데이터 1개당 다운로드 수
교통	1,293,925,897	73.9	4,792,318	보건	1,362,205	0.1	782
환경	443,692,130	25.3	1,308,826	도시관리	1,097,418	0.1	5,657
일반행정	4,755,752	0.3	7,822	산업경제	464,104	0.0	820
문화관광	3,423,982	0.2	5,358	복지	132,980	0.0	401
안전	2,857,902	0.2	27,747	교육	82,486	0.0	825
합계	1,751,794,856	100.0	358,094				

주 : 2017년 7월부터 2018년 8월까지 데이터별 다운로드 수 및 OpenApi 호출건수의 합계를 의미함
단위: 다운로드 수는 건, 비율은 %, 데이터 1개당 다운로드 수는 건

염 등 7개 영역에서 주로 활용되었다.

종합하여 살펴보면 공공데이터의 개방건수 증가와 공공빅데이터 활성화 간 상관관계는 낮은 것으로 나타났다. 데이터의 가용성이 증가함에 따라 교통·환경 등의 분야에서는 활용건수의 증가를 보였다. 그러나 공공의 정형데이터 위주의 개방이 이루어지면서 갱신주기가 길어지거나 표준화되지 않은 데이터들로 인해 데이터의 품질을 저하시킬 수 있고 연계성이 떨어질 수 있다. 따라서 반정형 혹은 비정형 데이터의 개방과 활성화를 위한 지속적인 노력을 해야한다. 또한 오픈Api의 서비스 비율이 높아지면서 교통환경 등 특정분야에서의 민간 활용성은 제고되고 있지만 더 폭 넓은 영역에서의 가치 있는 공공데이터 발굴이 필요하다.

4.2 서울시 빅데이터캠퍼스 현황과 이용실태

먼저 빅데이터캠퍼스에서 제공중인 데이터를 살펴보면 열린데이터광장에서 제공하는 데이터와 함께 공공·민간의 빅데이터 45건, 공간정보 473건을 공유한다. 공

공·민간의 빅데이터는 민간 14종과 공공 31종으로 주요 민간참여기관은 KB국민은행, KB카드, 나이스지니데이터, 신한카드, SKT, KT, 오픈메이트 등이 있었고 주로 소비, 소득, 인구/가구 등의 데이터를 제공하고 있었다. 이에 대한 자세한 내용은 <표 6>과 같다.

빅데이터캠퍼스에서는 시민들의 분석활동 지원 및

<표 6> 빅데이터 캠퍼스 제공 데이터

유형	구분	데이터명	적재주기	제공기관
민간	소비	장애인 복지카드 이용현황	1회	신한카드
민간	소비	3개축제카드이용DB	1회	신한카드
민간	소비	카드매출 데이터	월	KB카드
민간	시설	아파트 DB	연	오픈메이트
민간	시설	주요/집객시설	반기	오픈메이트
민간	인구/가구	유동인구 (월별 유동인구, 거주지)	월	SKT / KT
민간	인구/가구	유동인구 (일별 유동인구, 거주지)	월	SKT
민간	인구/가구	추정상주인구	반기	오픈메이트
민간	인구/가구	유동인구통계DB	1회	KT
민간	지역	블록 및 속성 데이터	연	오픈메이트
공공	교통	대중교통 환승데이터 (교통카드 거래내역)	월	서울시(스마트 카드사)
공공	교통	대중교통 이용통계	월	
공공	교통	도로별 속도통계	일	
공공	교통	교통시설(지하철, 버스정류장)	필요시	서울시
공공	교통	녹색교통운동	-	녹색교통
		자전거 이용 정보		운동본부
공공	교통	표준노드링크	반기	국토교통부
공공	교통	교통안전시설물	반기	서울시
공공	부동산	부동산실거래가	월	서울시
공공	부동산	전월세가	월	서울시
공공	부동산	공시지가	반기	서울시
공공	상가/기업	인허가업소DB	월	서울시
공공	상가/기업	사업체 통계 DB	년	통계청
공공	상가/기업	사업체 조사 DB(대시민용)	년	서울시
공공	소비	추정매출액(블록단위)	월	서울시
공공	시설	시설물(가판대, 구두수선대)	월	서울시
공공	시설	건축물대장	월	국토교통부

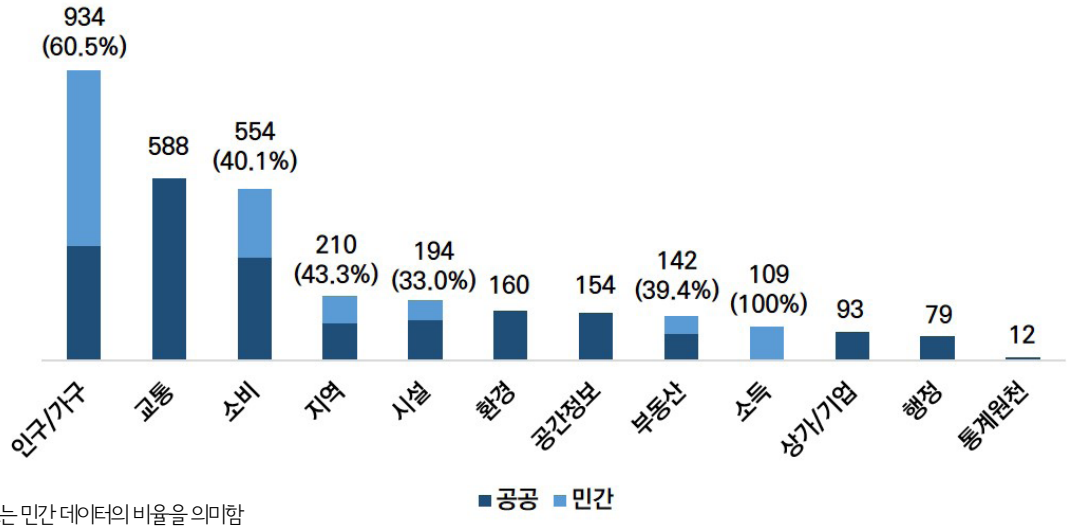
공공	시설	시설물(노인복지시설)	시점	서울시
공공	시설	건물DB	반기	서울시
공공	시설	도로명주소도로링크(도로구간)	반기	서울시
공공	인구/가구	길단위 추정 유동인구 (도로명주소 단위)	월	서울시
공공	인구/가구	거주인구(행정동)	월	행자부
공공	인구/가구	직장인구	반기	오픈메이트
공공	인구/가구	직장인구(집계구)	연1회	서울시
공공	인구/가구	인구센서스 통계	5년	통계청
공공	인구/가구	택시운행 통행량 (택시운행분석데이터)	월	서울시
공공	인구/가구	직업직종	년	통계청
공공	지역	골목상권영역 및 프로파일링	반기	서울시
공공	지역	핵심(발달)상권 영역	년	중소기업청
공공	환경	기상관측정보(기온, 강수 등)	비정기	서울시
공공	환경	강우량정보	일, 1분	서울시
공공	환경	대기환경정보	일	서울시
공공	인구/가구	생활인구	월	KT
민간	소비	신한카드 매출데이터	월	신한카드
민간	산업경제	기업 정보	연 1회	통계청

빅데이터 관련 지식공유를 위해 공모전 및 교육프로그램을 운영하고 있다. 데이터를 주제로 하는 자체 공모전을 3회 개최하였고 대학교 또는 공공기관과 연계하여 협력 컨퍼런스를 4회 진행하였다. 강의서비스는 빅데이터 강좌, 대학교 정규강의 및 기업교육과 연계해 실제 데이터를 온·오프라인 강의를 통해 지식공유를 추구하고 있다. 대학교 정규 강의 연계는 홍익대학교 도시공학과와 GIS 도시공간분석 수업을 연계하여 2회를 진행하였다. 또한, 일반인 및 입주자를 대상으로 데이터 분석 및 케이스 스터디 형식으로 오픈강의를 진행하였다. 보험비즈니스 사례(1회), 개인정보 비식별화 실습교육(1회), 서울시 축제 분석 따라하기(2회) 및 유동인구 데이터 정제·분석·시각화 교육(3회)로 총 7회 진행이 되었다. 이 밖에 기업 교육 연계 교육과정을 추진중이다.

빅데이터캠퍼스의 데이터 이용 추이를 살펴보면 2018년 11월까지 732건 입주가 신청되었고 계속해서 증가하는 추세를 보였다. 2017년 1월 112건과 비교하였을 때 약 6.5배 증가한 것으로 확인됐다. 한편 732건의 입주 중 대학교가 537건(73.4%), 민간 105건(14.3%), 공공기관 32건(4.4%) 순으로 높게 나타나 상대적으로 교육 및 연구목적으로 많이 활용되고 있는 것으로 추측할 수 있다.

데이터 유형별 이용건수를 살펴보면 <그림 3>에서 확인할 수 있듯이 인구/가구가 934건으로 가장 많고 다음으로는 교통(588건), 소비(554건), 지역(210건), 시설(194건) 등의 순으로 나타났다. 이에 대한 민간 데이터의 제공 비율이 높은 분야는 인구/가구(60.5%), 소비(40.1%), 지역(43.3%), 부동산(39.4%), 소득(100%) 등으로 나타났다.

가장 많이 활용이 된 데이터 상위 10개를 살펴보면



주 : 괄호는 민간 데이터의 비율을 의미함

<그림 3> 데이터 분야별 이용 건수

<표 7>과 같다. SKT의 유동인구 데이터가 15.8%로 가장 높았으며, 다음으로 서울시의 대중교통 환승데이터(8.4%), 서울시의 추정매출액(7.5%) 순으로 나타났다. 이중 민간 데이터는 유동인구, 카드매출, 소득수준 그리고 블록별 속성데이터가 포함되었다.

분석에 대한 결과를 살펴보면 빅데이터캠퍼스는 빅데이터 45건과 공간정보 473건을 공유하고 있으며 누적이용자 수는 증가 추세에 있다. 입주 유형은 대학교의 비율이 약 73%로 교육 및 연구 목적으로의 활용도가 높은 것으로 판단할 수 있다. 그러나 민간 데이터이

<표 7> 이용건수 상위 10개 데이터

데이터명	유형	데이터제공기관	이용건수(비율)
유동인구(일별 및 월별)	인구/가구	SKT	511 (15.8%)
대중교통 환승데이터(교통카드 거래내역)	교통	서울시	271 (8.4%)
추정매출액(블록단위)	소비	서울시	241 (7.5%)
SDW 보유 공간정보, 정책지도, 수치지도, 공간정보플랫폼 자료	공간정보	서울시	154 (4.8%)
카드매출 데이터	소비	KB카드	145 (4.5%)
소득수준데이터	소득	나이스지니데이터	109 (3.4%)
길단위 추정 유동인구 (도로명주소 도로링크 단위)	인구/가구	서울시	96 (3.0%)
거주인구(행정동, 법정동)	인구/가구	행정자치부	91 (2.8%)
블록 속성 데이터	지역	오픈메이트	91 (2.8%)
골목상권영역 및 프로파일링	지역	서울시	88 (2.7%)

출처 : 서울시 내부자료(18년 11월 기준)

용 비율은 상대적으로 떨어지는 것으로 나타났다. 캠퍼스 운영의 효율성을 높이기 위해서는 온라인 플랫폼을 발전시켜 활성화 전략을 추진할 필요가 있다. 또한 데이터 공유를 위한 공공·민간의 참여를 촉진할 방안을 모색하고 공모전·교육프로그램을 확대해야 한다.

5. 토론

본 연구는 빅데이터의 활용방안 관점에서 서울시의 빅데이터 관련 정책을 살펴보고 서울시가 운영하는 열린데이터광장과 빅데이터캠퍼스의 이용실태를 파악함으로써 공공빅데이터 활성화를 위한 정책 제언을 하는데 그 목적을 두고 있다. 지난 몇 년간 데이터의 개방건수는 지속적으로 늘어났으나 이러한 과정 속에서 데이터를 활용한 사업화나 민간영역에서의 적용사례는 아직 미흡한 실정이다. 또한 데이터의 생애주기에 따라 그 특성과 활용방안이 다를 수 있음에도 특정영역 안에 국한되어 실질적인 데이터의 적용 성과는 낮은 상황이다. 따라서 서울시의 빅데이터 관련 정책 혹은 운영되는 서비스의 사례들의 실태 파악을 통해 단점 및 장점을 정리하고 이를 토대로 개선점과 공공데이터 활성화를 위한 정책 제언에 대한 논의를 하고자 한다.

구체적으로 데이터의 생애주기와 각 단계별 공공데이터의 제공자와 이용자 측면에서 공공빅데이터 활성화를 위한 방안을 제시하고자 한다. 먼저, 데이터의 수집 및 저장관리 단계의 제공자 측면에서는 거버넌스, 전문인력 양성, 제도 개선 등 외부 환경변화가 활성화를 유도한다. 예를 들어 데이터의 공유문제는 '공공-민간'보다 '공공-공공' 간 공유체계 수립이 중요하다고 할 수 있다. 이용자 측면에서는 데이터 수집과 개방에 대한 적극적인 요구와 함께 그들이 데이터 수집·저장관리 단계에 참여할 수 있는 오픈플랫폼 구축이 필요하다. 개인정보 등과 같은 이슈에 대한 적절한 방안을 정립한다면 이용자 역시 데이터를 생성하는 주체의 역할을 수행할 수 있다. 분석에서는 주로 현재 데이터의 표준화 혹

<표 8> 데이터 생애주기/제공자 및 사용자 유형별 활성화 촉진 및 저해요인

생애주기/유형	제공자 측면	이용자 측면
수집· 저장관리	A 공공데이터 개방(+) A 내부 전문 인력 양성(+) A 외부 전문 인력의 충원(+) A 공공분야 데이터 전문 중개·지원기관 (+) A 조직 내 낮은 문화·인식(-) A 거버넌스의 부재(-)	A 개인정보 보호 및 활용 양립방안(+) A 데이터 공유·거래 거버넌스의 구축(+) A 미공개 빅데이터에 대한 적극적 요구(+) A 시민참여형 오픈플랫폼 구축(+)
분석	A 데이터 형식 표준화 및 가용성 등 전문적인 품질관리(+) B 빅데이터캠퍼스등 인프라 구축(+) B OpenApi 등 개발자를 위한 서비스 제공비율의 증가(+) B 정형데이터 위주 개방(-)	A 데이터 신뢰도 향상(+) A 데이터 형식의 비일관성(-) A 데이터의 일회성 분석 사용(-) B 데이터의 낮은 갱신주기(-) B 교통·환경 등 특정분야 데이터 집중(-)
활용	A 공공기관 실적공개(+) A 빅데이터 분석을 위한 제도적 근거(+) B 분석사업 지원을 위한 재정 유동성(+) B 데이터에 대한 현장의 낮은 인식(-)	A 사용자 유형별 교육프로그램 구성(+) A 필요성에 대한 사회적 수용성(+) A 빅데이터 지식공유 아카이빙(+) A 데이터 서비스 체계 개선(+)

주 : 괄호의 +는 촉진요인을, -는 저해요인을 의미하고 A는 선행연구, B는 실태조사를 의미함
출처 : 성욱준(2016 2017), 전병진 외(2017), 본 연구결과의 내용을 바탕으로 작성하였음

은 외부 데이터와의 연계와 같은 데이터의 품질문제가 대두되었다. 즉, 공공데이터의 양적인 개방보다는 질적인 개선이 먼저 선행되어야 한다. 이를 위해 빅데이터캠퍼스 등 물리적 인프라도 중요하지만 데이터의 품질관리를 위한 개선체계 마련과 지속적인 관심을 유도할 필요가 있다(김정선 등 2014). 활용에서는 생산자·소비자의 데이터에 대한 인식개선 및 제고, 이용자 유형별 데이터 서비스 체계 구성 등 데이터에 대한 접근성 향상, 기술·정책 등 지식을 공유하는 아카이빙이 필요하다. 특히, 제공자 측면에서 본다면 데이터 활용을 위한 제도적 기반마련과 예산의 유동성을 통해 적시적소에 예산을 투입해 공공데이터의 활용도를 높일 필요가 있다.

분석결과와 해석을 토대로 다음과 같이 서울시의 공공빅데이터 활성화를 위한 기본 방향을 제시하였다. 첫째, 공공데이터의 활성화와 적용을 위해선 데이터의 양보다는 질에 집중해야 하며 앞으로의 정책적 목표는 시장수요가 많고 활용가치가 높을 것으로 예상되는 빅데이터를 생성하고 발굴하는 것이다. 둘째, 데이터에 대한 권리가 조직 간 분산된 구조를 유지하고 있어 개인 혹은 필요한 집단으로의 연계가 어렵고, 이를 중개해 줄 정보화 조직의 위상이 낮기 때문에 데이터의 생성과 수집 그리고 관리할 수 있는 권한을 기반으로 거버넌스를 재구축 해야한다. 셋째, 데이터를 다루는 사람 혹은 조직 등 관련 이해관계자들의 데이터에 대한 인식 개선과 태도의 변화가 요구되며, 시민·공무원·민간 역시 데이터의 중요성을 인지할 수 있도록 사회전반에 걸친 개념 확립이 필요하다. 넷째, 개인 및 집단의 참여를 장려하는 플랫폼을 통해 데이터에 대한 권리와 정보 주체에 대한 실질적 이익이 보장되어야 한다. 마지막으로 이용자 및 데이터의 형태에 따라 데이터의 서비스 활용을 다르게 적용하여 접근성을 높여야 한다. 어떻게 서비스를 할지에 대한 전략으로부터 데이터의 활용이 연결될 것이다.

6. 결론

본 연구는 빅데이터와 직접적으로 연결되는 정책 소개와 서울시가 주도하는 빅데이터 활성화 프로그램에 대한 실태조사를 기반으로 실질적인 서울시의 공공빅데이터 활성화 방안의 방향성을 제시하였다. 이에 본 연구는 다음과 같은 학술적 및 실무적 시사점을 가진다.

지금까지 여러분야에서 빅데이터의 소개 및 도입, 활용방안 등에 대하여 많은 연구가 이루어졌다. 그러나 이러한 연구들은 특정 산업분야에 초점을 맞추거나 전반적인 동향과 발전방향에 대하여 논의해왔다. 본 연구는 서울이라는 특정 지역안에서 실질적으로 운영되고 있는 현상에 대하여 실태파악을 위한 조사를 실시하였고, 이를 정리함으로써 기존 정책으로 연결될 수 있는 빅데이터 연구를 확장하였다. 또한, 이용실태 분석에 있어 데이터의 생애주기 틀을 적용하여 구분하고 각각에 해당하는 분석 결과를 제시했다는 이론적 시사점을 가진다.

실무적으로는 기존 서울시의 빅데이터 관련 정책 및 활동을 살펴보고, 이중 특정 프로젝트에 대해 집중 조명하였다. 이를 토대로 현재 잘 이루어지는 부분 혹은 충분한 데이터가 제공되고 있는 영역 등을 확인함과 동시에 충분한 잠재력을 가지고 있음에도 불구하고 활용도가 떨어지는 부분을 파악하여 운영시설 및 공공빅데이터 활용에 대한 발전방향을 제시하였다. 더불어 서울시의 정책 수립과정에 직간접적으로 영향을 미칠 수 있는 방향성을 제시하였다. 따라서 본 연구는 이후 서울시의 공공빅데이터 관련 정책 수립 및 운영 전략을 세움에 있어 기초자료로 활용될 수 있을 것이다.

위와 같은 시사점에도 불구하고 본 연구는 현재 시행중인 정책 혹은 프로그램의 실태조사를 기반으로 연구를 진행하였기 때문에 실질적인 빅데이터 관련 이해관계자들의 인식이나 행동의도 등을 고려하지 못했다는 한계점을 가진다. 따라서 이후 연구에서는 빅데이터와

직간접적으로 연결될 수 있는 당사자들의 의견을 설문조사나 인터뷰 등을 통해 실증분석하여 정책적 제언에 도움이 될 수 있는 의미를 추출할 필요가 있다.

참 고 문 헌

[국내 문헌]

1. 강진동 2016. "서울시 빅데이터 활용," *교통기술과 정책* (13), pp. 37-42.
2. 공공데이터전략위원회 2016. 제2차 공공데이터 기본계획.
3. 관계부처 합동 2013. 정부 3.0 추진 기본계획.
4. 관계부처 합동 2018. 공공데이터 혁신전략.
5. 권영일, 김진철, 김성현 2017. "빅데이터 정책의 현황과 미래," *한국지능정보시스템학회 학술대회논문집*, pp. 90-91.
6. 김구 2017. "OECD 국가의 공공데이터 개방에 관한 연구," *한국공공관리학보* (31:2), pp. 201-225.
7. 김도우 2017. "공공안전 분야의 빅데이터 정책과 연구동향," *한국행정학회 학술발표논문집*, pp. 3044-3069.
8. 김동성, 김종우, 이홍주, 강만수 2017. "공공부문 데이터의 경제적 가치평가 연구: 소상공인 신용보증 데이터 사례," *지식경영연구* (18:1), pp. 67-81.
9. 김동완 2013. "빅데이터의 분야별 활용사례," *경영논총* (34), pp. 39-52.
10. 김민길, 유정호, 조민효 2018. "행복 (공공) 기숙사 도입이 대학가의 임대료에 미친 영향에 관한 연구-서울시 빅데이터 캠퍼스를 활용하여," *한국정책학회보* (27:4), pp. 35-67.
11. 김성태 2011. "신가치창출 엔진, 빅 데이터의 새로운 가능성과 대응 전략," *한국정보화진흥원 IT & Future Strategy* (18), pp. 1-34.
12. 김정선, 권은주, 송태민 2014. "분석지의 확장을 위한 소셜 빅데이터 활용연구-국내," *지식경영연구* (15:3), pp. 169-188.
13. 박민정, 채상미 2017. "빅데이터 환경 형성에 따른 데이터 감시 위협과 온라인 프라이버시 보호 활동의 관계에 대한 연구," *지식경영연구* (18:3), pp. 63-80.
14. 배동민, 박현수, 오기환 2013. "빅데이터 동향 및 정책 시사점," *정보통신방송정책* (555), pp. 37-74.
15. 서울특별시 빅데이터캠퍼스(https://bigdata.seoul.go.kr)
16. 서울특별시 2015. 서울디지털 기본계획2020.
17. 서울특별시 2018. 서울특별시 정보화 시행계획.
18. 서울특별시 2018. 서울시 빅데이터 시행계획.
19. 성욱준 2016. "공공부문 빅데이터 정책 활성화 연구," *한국정책학회보* (25:2), pp. 125-149.
20. 성욱준 2017. "데이터 생애주기 관점에서 본 공공부문 빅데이터 활성화 방안," *한국지역정보화학회지*, (20:2), pp. 25-41.
21. 신동희, 김용문 2015. "국내 재난관리 분야의 빅데이터 활용 정책방안," *한국콘텐츠학회논문지*, (15:2), pp. 377-392.
22. 오재인 2017. "마이더스의 손, 빅데이터 시대를 맞이하여: 지자체의 공공정책을 위한 빅데이터 전략과 활용 방안," *지역정보화* (104), pp. 38-41.
23. 유영성, 빈미영, 옥진아, 최조순, 천영석 2014. "지자체의 공공 빅데이터 정책 사례연구," *정책연구*, pp. 1-50.
24. 이각범 2011. *빅데이터를 활용한 스마트 정부 구현 (안)*, 국가정보화전략위원회 보고서.
25. 이강용, 남궁현, 심재철, 조기성, 류원 2012. "공공분야에서의 빅 데이터 활용을 위한 지식자산

(Knowledge Base) 구축," *정보과학회지* (30:6), pp. 40-46.

26. 이창범 2013. "개인정보보호법제 관점에서 본 빅데이터의 활용과 보호 방안," *법학논총* (37:1), pp. 509-559.
27. 임용재, 백선경, 연승준 2012. "빅데이터 시대의 경쟁력 확보를 위한 선택과 집중," *한국통신학회지* (정보와통신) (29:11), pp. 3-10.
28. 임은선, 이영주, 황명화, 오창화 2018. "빅데이터 시대의 국토정책 추진방향과 과제," *국토정책 Brief*, pp. 1-8.
29. 전병진, 김희웅, 2017. "공공 빅데이터 개방 및 활용 활성화 방안에 대한 연구," *정보화정책* (24:3), pp. 27-41.
30. 정보통신정책연구원 2017. 4차 산업혁명 시대의 데이터 경제 활성화 전략.
31. 정지형, 김강훈 2012. "한국과 미국의 공공부문 빅데이터 활용 현황 분석," 한국정보과학회 학술발표 논문집 (39:2D), pp. 243-245.
32. 조성우 2011. "빅데이터 시대의 기술," KT 종합기술원, pp. 5-7.
33. 조영임 2013. "빅데이터의 이해와 주요 이슈들," *한국지역정보화학회지* (16:3). pp. 43-65.
34. 토지주택연구원 2015. 토지주택분야 적용을 위한 공간 빅데이터 활용 기초연구.
35. 표순희, 김윤형, 김혜선, 김완종 2015. "도서관 빅데이터 서비스 모형 개발에 관한 연구: 공공도서관을 중심으로," *정보관리학회지* (32:2), pp. 63-86.
36. 한국데이터진흥원 2018. 2017년 데이터 산업 현황 조사.
37. 한국정보통신기술협회 2015. 데이터 생애주기 기반 빅데이터 도입 및 활용지침.

[국외 문헌]

1. Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., and Zhao, J. L. 2016. "Transformational Issues of Big Data and Analytics in Networked Business," *MIS quarterly* (40:4). pp. 807-818.
2. Chen, C. P., and Zhang, C. Y. 2014. "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Information sciences* (275), pp. 314-347.
3. Fredriksson, C., Mubarak, F., Tuohimaa, M., and Zhan, M. 2017. "Big data in the public sector: A systematic literature review," *Scandinavian Journal of Public Administration* (21:3), pp. 39-62.
4. Gurin, J. 2014. "Big data and open data: How open will the future be," *ISJLP* (10:3), pp. 691-704.
5. IDC. 2015. Digital Transformation(DX): An Opportunity and an Imperative.
6. Ko, J. G., Bae, Y. S., Park, J. Y., and Park, K. 2014. "Technologies Trends in Image Big Data Analysis," *Electronics and Telecommunications Trends* (29:4), pp. 21-29.
7. Laney, D. 2001. "3D data management: Controlling data volume, velocity and variety," *META group research note* (6:70), pp.1
8. McKinsey, 2011, The next frontier for innovation, competition, and productivity, McKinsey Global Institute Report.
9. McKinsey, 2018, Notes The AI Frontier Insights from Hundreds of Use Cases. McKinsey Global Institute Report.

저 자 소 개

최 봉 (Bong Choi)

현재 서울연구원 시민경제연구실 선임연구위원으로 재직 중이다. 성균관대학교 산업공학과(1983)를 졸업하였고, 한국과학기술원(KAIST) 산업공학과에서 석사(1987), 성균관대학교 경영학과에서 경영정보시스템 전공으로 박사(2005)학위를 취득하였다. 주요 관심분야는 지능형 정보기술, ICT 산업정책, 혁신전략 등이다. 지금까지 Total Quality Management and Business Excellence, International Journal of Operations and Production Management 등 주요 학술지에 논문을 게재하였다.

윤 종 진 (Jongjin Yun)

현재 서울연구원 시민경제연구실 연구원으로 재직 중이다. 서울시립대학교에서 도시공학 석사학위를 취득하였다. 주요 관심분야는 지역경제, 스마트시티, 혁신성장 등이다.

엄 태 휘 (Taehyee Um)

현재 경희대학교 호텔관광대학 일반대학원 박사과정으로 재학 중이다. 경희대학교에서 관광학 석사학위를 취득하였고, 스마트관광연구소에서 연구원 및 호텔관광대학 문화관광산업학과 조교, 기업경영학회 편집간사직을 맡고 있다. 주요 관심분야는 호텔경영, 스마트관광, 온라인여행사, 소비자관광행동 등이다. 지금까지 한국관광학회, 서비스경영학회, 한국호텔외식관광경영학회, ENTER, PACIS 등 주요 학술지에서 논문을 발표하였다.