

요약

교통 등 5대 핵심분야 공공빅데이터 발굴하고 품질관리체계 구축·플랫폼 간 협업관계도 형성

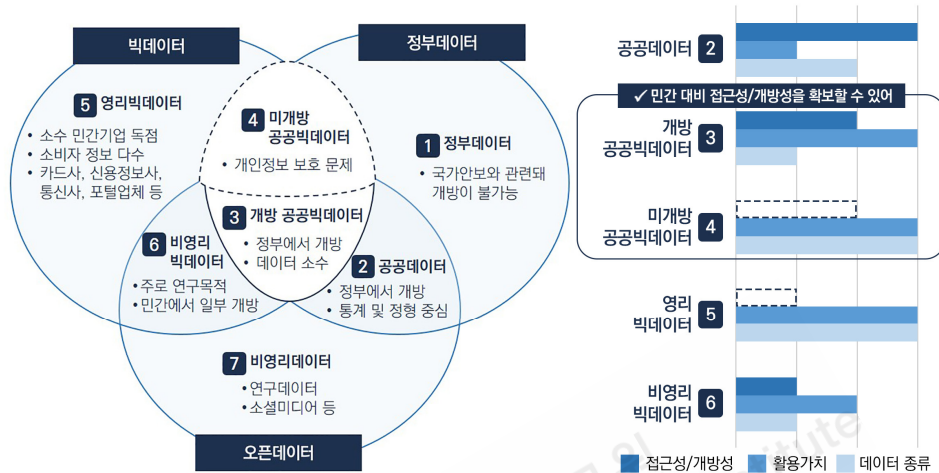
빅데이터, 4차 산업혁명·데이터경제시대 혁신자원으로 '중요성 커져'

빅데이터에 대한 관심이 뜨겁다. 빅데이터는 폭발적인 데이터의 양적 증가와 함께 주목받기 시작하였지만, 지금은 4차 산업혁명과 데이터경제 등 새로운 패러다임의 핵심 기술 또는 자원으로 인식되면서 중요성이 더욱 커지고 있다. 4차 산업혁명은 데이터의 생성과 이동에 따른 핵심기술(사물인터넷, 클라우드, 빅데이터, 인공지능 로봇 등) 간 융합에서, 데이터경제는 생성부터 분석까지 데이터가 생산요소로 활용되어 새로운 편익을 늘린다는 점에서 빅데이터를 근간으로 삼고 있다. 두 논의의 관점은 다를지라도 새로운 가치를 창출하기 위해 '어떻게 양질의 데이터를 생성하고 잘 유통하여 활용할 것인가?'라는 질문은 유사하다. 공공·민간 모두에게 데이터는 기존 모델의 디지털 전환을 위한 필수요소이자 혁신 자원으로 인식되고 있다. 따라서 양질의 빅데이터를 확보할 수 있는지 여부가 공공서비스의 혁신과 기업의 성장가능성을 결정하게 될 것이다.

공공빅데이터, 기업·시민이 쉽게 활용 가능 '디지털전환 파급효과 커'

공공빅데이터에 대한 명확한 이론적 정의는 없다. 다만, 본 연구에서는 정부데이터, 오픈데이터, 빅데이터에 따라 유형을 구분하였다. 공공빅데이터는 '공공부문이 가진 빅데이터'로 볼 수 있으며 개방여부에 따라 구분된다. 미개방 공공빅데이터는 의료·금융·통신·위치정보 등 활용성이 높을 것으로 예상되나, 개인정보보호 등 제도적 여건으로 인해 개방이 이뤄지지 않은 상태다. 개방 공공빅데이터는 개방되어 있으나 교통·환경 등 데이터 종류가 소수에 그치고 있다. 영리빅데이터는 통신사·카드사·신용정보사·포털업체 등 소수의 민간기업이 정보를 독점하고 있어 활용성은 높다고 인지되나 접근/개방성이 취약한 상황이다. 앞으로도 정보를 독점한 기업들

은 데이터를 자산으로 인식하며 개방에 적극적이지 않을 것으로 예상된다. 공공빅데이터는 민간 수준의 활용가치를 담보하면서도 일반기업과 시민이 쉽게 접근해 활용할 수 있기 때문에 디지털 전환에 미치는 파급효과는 크다고 볼 수 있다.



출처 : 연구진 작성

[그림 1] 빅데이터, 오픈데이터, 정부데이터에 따른 유형과 특징

서울시, 빅데이터캠퍼스 운영·전담조직 신설로 이용 활성화 지속 노력

국내에서는 2013년 이후 적극적으로 공공데이터가 개방되었다. 우리나라는 2015년 OECD 공공데이터 개방지수 1위를 차지하는 등 높은 성과를 이뤘지만, 민간에서 요구하는 양질의 데이터가 부족하고 품질문제가 대두됨에 따라 민간의 이용은 저조하다. 특히, 그간 공공데이터를 개방하는 정책과 빅데이터 산업을 활성화하려는 정책 간 거버넌스가 구축되지 않아 데이터의 생성과 활용이 유기적으로 연결되기 어려웠다.

서울시는 2013년 빅데이터를 활용해 심야버스(올빼미버스) 노선을 선정하는 등 선도적으로 공공정책을 수립했다. 정보인프라의 구축·유지와 함께 새로운 IT패러다임에 대응하기에는 지방 정부의 정보화 조직 위상이 낮고 예산과 인력이 부족한 것이 현실이지만, 빅데이터를 중심으로 조직을 개편하고 전담조직을 빠르게 신설하였다(활용 - 서울디지털재단, 생성 및 수집관리 - 통계데이터담당관). 또한, 상암 빅데이터캠퍼스 및 개포디지털혁신파크의 도시데이터사이언스

연구소 등 물리적인 인프라를 갖췄기에 개인정보 등 제도적 여건이 마련되지 않은 상황에서도 시민·민간에게 공공빅데이터를 개방하고 활용을 촉진하려는 노력을 꾸준히 견지할 수 있었다. 2018년에는 「서울특별시 데이터의 제공 및 이용 활성화에 관한 조례」를 제정해 민·관 빅데이터 심의위원회를 구성하고 빅데이터 시행계획을 수립해 빅데이터 정책을 체계적으로 추진하고 있다. 하지만 아직까지 데이터 생성은 각 부처에서 이뤄지고, 수집·저장관리는 통계데이터담당관에서, 이용 활성화는 서울디지털재단에서 맡는 등 데이터 생애주기 단계별로 조직이 구분되어 있는 상황이다. 데이터의 생성과 수집 그리고 관리할 수 있는 권한을 바탕으로 협력적 거버넌스가 재구성되어야 할 시점이다.

세계 주요도시, 시민참여 플랫폼·IoT 오픈생태계·혁신클러스터 구축

해외 주요도시(시카고·뉴욕·베이징·바르셀로나·싱가포르)의 빅데이터 활성화를 위한 정책사례를 살펴보았다. 시카고는 공공데이터의 통합 관리와 사물인터넷을 활용한 오픈플랫폼(Array of Things, AoT)을 구축했다. AoT는 도시 전역의 온도, 기압, 오염물질, 소음 등 주변 환경을 실시간으로 측정해 빅데이터를 생성한다. 또한 설계도면 개방을 통해 민간참여를 피하고 있다.

뉴욕은 데이터 개방 법률을 제정하고 정보·통신기술 전략계획을 수립하는 등 제도적 여건을 마련했다. 특히, 시 정부가 주도적으로 연구·교육기관, 스타트업, 비영리단체, 시민단체 등 다양한 이해당사자와 거버넌스를 구축해 데이터 생태계를 형성하고 있다.

베이징은 ‘중관춘’ 혁신클러스터의 조성 and 빅데이터산업원의 설립을 통해 IoT 및 클라우드컴퓨팅의 융합발전을 촉진하고 있다. 이에 따라 중관춘은 4차 산업 관련 기업의 창업 및 성장을 유도해 세계에서 두 번째로 유니콘기업이 많은 스타트업 메카로 부상할 수 있었다. 또한 통신·금융 민간기업과 협력해 차량공유서비스, 교통정보플랫폼 등 도시문제 해결을 모색 중이다.

바르셀로나는 도시 전역에 센서를 설치하고 개방형 IoT 센서 플랫폼을 구축 중이다. 모든 센서는 통합 관리가 가능하며 기술적 솔루션을 민간에게 제공해 참여를 독려하고 있다. 또한, 시민참여 플랫폼은 빅데이터 프로젝트에 대한 시민의 의견을 수렴하고 시민이 직접 참여할 수 있어 개인정보보호·교육 등 시민의 디지털 주권 회복을 도모하고 있다.

싱가포르는 데이터 수집·분석 개발을 위해 해외 연구기관 및 기업과의 협력관계를 구축하고

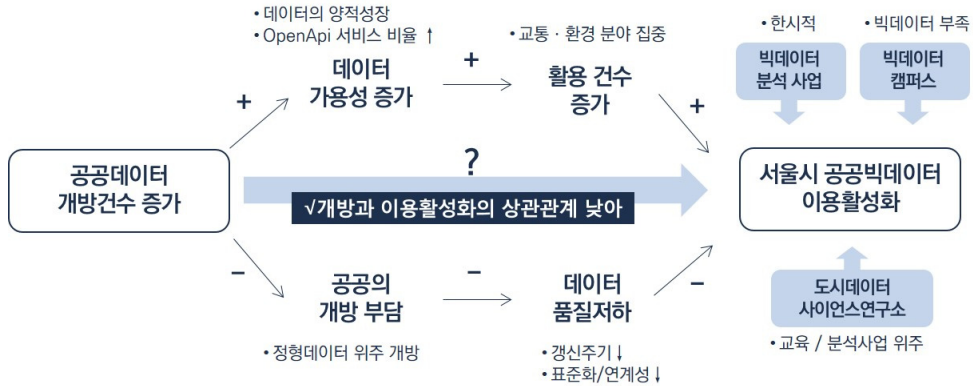
도시국가 차원의 프로젝트(3D 가상플랫폼, 국가위험관리시스템 등)를 진행 중이다.

설문조사 결과 “교통환경 수요가 많아”...빅데이터다운 데이터는 부족

서울시 공공빅데이터의 실태를 확인하기 위해 서울열린데이터광장과 빅데이터캠퍼스를 살펴보고 서울 소재 개발자·스타트업을 대상으로 설문조사를 실시하였다. 열린데이터광장은 2013년 대비 약 7.2배 증가한 4,892건(2018년 8월)의 공공데이터를 개방했다. 또한 보건·문화관광·일반행정·산업경제 분야 순으로 개방이 많았고, OpenApi 서비스 비율이 약 90%에 달하는 등 데이터의 가용성은 증가했다. 하지만 정형데이터 위주로 개방했고, 전체 약 9%만이 한 달 이내로 정기적인 업데이트가 이뤄지고 있다. 특히, 교통·환경 분야는 개방 건수는 적지만 이용건수가 전체 약 99%에 달해 수요가 높은 분야의 개방이 잘 이뤄지지 않는 것으로 보인다. 또한 관측치 10,000개 이상의 데이터는 많지 않아 빅데이터다운 데이터는 부족한 실정이다.

빅데이터캠퍼스는 빅데이터 45건과 공간정보 473건을 추가로 시민·민간과 공유 중이며, 2016년부터 누적이용자수도 증가 추세다. 입주자 유형은 교육 및 연구 목적으로 이용하는 대학교가 약 73%로 가장 높다. 유형별 데이터 이용건수는 인구/가구·교통·소비 등의 순으로 높고, 민간데이터 이용 비율(약 34%)은 절반에 못 미치고 있다. 이용건수가 높은 데이터는 SKT 유동인구·서울시 대중교통환승데이터·서울시 추정매출액 순으로 나타났다.

서울시 소재 개발자·스타트업 설문조사 결과 전체의 약 47%가 서울시 데이터를 이용한 경험이 있었다. 절반 이상의 응답자가 공공데이터를 활용해 사업모델을 구상 중이었으며, 광고 및 사 용료·정기구독료 형태의 수익창출모델이 대다수다. 정책수요는 공공데이터 확대개방, 품질관리 강화, 주요 빅데이터 중점관리 순으로 응답률이 높게 나타났다.



[그림 2] 공공데이터 개방건수 증가와 활성화의 관계

데이터 생애주기별 이용 활성화 촉진·저해요인 고려해 정책 설계해야

국내 선행연구와 실태조사 결과를 바탕으로 데이터 생애주기 단계별 및 제공자/이용자별 활성화 촉진 및 저해 요인을 제시하였다. 수집·저장관리 단계의 제공자는 거버넌스, 전문인력 양성, 제도 개선 등 외부 환경변화가 요구된다. 반면 이용자는 개인정보 보호 및 활용 양립방안 구축, 적극적인 개방 요구, 데이터 수집·저장관리 단계에 참여할 수 있는 오픈플랫폼 구축이 필요하다. 분석단계에서는 주요 문제점으로 인지되는 데이터 품질문제가 주를 이루 양적 개방보다는 질적 개방으로 변화할 필요가 있다. 활용단계에서는 데이터 생산자·소비자의 인식 개선 및 제도, 이용자 유형에 따른 서비스체계 구축을 통한 데이터 접근성 향상, 기술·정책을 공유할 수 있는 아카이빙이 필요하다. 즉, 시민·개발자·연구자 등 이용자 유형에 따라 ‘데이터를 쉽게 찾아 다운 받고 분석할 수 있는 환경을 조성하는 것’이야말로 활성화의 본질이라 볼 수 있다.

[표 1] 데이터 생애주기/제공자-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 총원 (+) A 공공분야 데이터 전문 중개지원기관 (+) A 조직 내 낮은 문화인식 (-) B 빅데이터 저장소 기반 거버넌스 (+)	A 개인정보 보호 및 활용 양립방안 (+) A 데이터 공유-거래 거버넌스의 구축 (+) A 미공개 빅데이터에 대한 적극적 요구 (+) B 시민참여형 오픈데이터 플랫폼 (+) B 사물인터넷 오픈생태계 구축 (+)
분석	A 데이터 형식 표준화 및 가용성 등 전문적인 품질관리 (+) C 빅데이터캠퍼스 등 인프라 구축 (+) C OpenApi 등 개발자를 위한 서비스 제공 비율의 증가 (+) C 정형데이터 위주 개방(-)	A 데이터 신뢰도 향상 (+) A 데이터 형식의 비밀관성 (-) A 데이터의 일회성 분석 사용 (-) B 분석 프로젝트 기반 데이터 발굴유통 (+) C 데이터의 낮은 갱신주기 (-) C 교통·환경 등 특정분야 데이터 집중 (-)
활용	A 공공기관 실적공개 (+) A 빅데이터 분석을 위한 제도적 근거 (+) C 분석사업 지원을 위한 재정 유동성 (+) C 데이터에 대한 현장의 낮은 인식 (-)	A 사용자 유형별 교육프로그램 구성 (+) A 필요성에 대한 사회적 수용성 (+) A 빅데이터 지식공유 아카이빙 (+) A 데이터 서비스 체계 개선 (+) B 이용자 유형별 데이터 서비스 체계 (+)

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

‘서울공공빅데이터’ 발굴해 중점관리하고 협력 거버넌스도 구축 필요

연구결과를 바탕으로 서울시 공공빅데이터 활성화를 위한 기본방향을 제시하였다. 첫째, 공공 데이터를 바라보는 관점은 ‘양’에서 ‘질’로 변해야 하며, 개방이 아닌 생성을 통한 발굴로 정책 목표가 바뀌어야 한다. 둘째, 활성화를 저해하는 것은 데이터가 아닌 사람과 조직의 문제로 각 이해당사자의 인식의 변화와 개선이 필요하며, 데이터를 생성, 수집 그리고 관리할 수 있는 권한을 바탕으로 거버넌스를 재구축해야 한다. 셋째, 시민참여를 장려하는 온라인 플랫폼을 구축해 데이터에 대한 권리를 보장하고, 이용자 및 데이터 형태에 따라 서비스 방법을 달리해 데이터 접근성을 높여야 한다. 넷째, 사물인터넷·블록체인·인공지능 등 신기술을 통한 산업 육성과 공공빅데이터 생성 및 활용은 서로 공생 관계이므로, 서울시는 분야별 오픈플랫폼 생태계

를 구축하고 ‘플랫폼의 플랫폼’ 전략으로 플랫폼 간 협업관계를 형성해야 한다.

주요 세부과제로는 교통·환경·에너지·재정/금융·산업인터넷 등 5대 핵심 분야에 대해 공공빅데이터를 새롭게 발굴하고, 신산업성장의 기반이 되며 시민의 삶의 질을 높일 수 있는 가치 ‘서울 공공빅데이터’를 새롭게 지정해 중점적으로 품질을 관리해야 한다. 또한, 데이터 서비스와 공공·민간 소통창구 역할을 전담할 정보화진흥 실행조직을 신설하고, 기술조직 간 협업체계를 구축해야 한다. 구로(IoT)·양재(인공지능) 빅데이터캠퍼스를 신설하고 데이터 품질·표준·플랫폼 관련 R&D를 진행해 데이터 기반의 신산업 육성을 위한 기반을 마련해야 한다.

분야	신규 발굴사업	기본방향	
서울시 공공빅데이터 활성화	공공빅데이터 기반조성	<ul style="list-style-type: none"> • ‘서울공공빅데이터’ 선정 및 중점관리 • 5대 핵심 분야 공공빅데이터 발굴 • 데이터 품질관리체계 구축 및 관리 	1 1 1
	거버넌스·인프라	<ul style="list-style-type: none"> • 서울디지털재단 내 ‘정보화진흥본부’ 설치 및 운영 • 데이터 생애주기 기반 ‘서울시 빅데이터 디지털전환 거버넌스’ 구성 • 서울시민 참여 서울형 ‘Mydata’ 공통 플랫폼 • 서울연구데이터(SeTTA) 및 서울공간정보 플랫폼 구축 	2 1, 2, 3, 4 3 1, 3
	교육	<ul style="list-style-type: none"> • 서울시립대 도시과학대학원 내 도시빅데이터 석사과정 설립 (공무원 대상) • 서울시 및 산하기관 전 직원 데이터 관련 의무교육 실시 	2 2
	시범사업	<ul style="list-style-type: none"> • 시장창출형 빅데이터 활용 시범사업 확대 운영 (IoT 및 인공지능 연계) • 비정형 공공빅데이터를 활용한 도시혁신 시범사업 	1 1
	R&D	<ul style="list-style-type: none"> • 서울시 IoT 및 빅데이터 설계, 품질, 연계, 표준화, 플랫폼 등 정보화 연구 	4
	산업진흥	<ul style="list-style-type: none"> • IoT(구로) 및 인공지능(양재) 빅데이터캠퍼스 신설, 센터 간 역할 정립 	4
	제도개선	<ul style="list-style-type: none"> • 서울시 물품·용역사업의 데이터 관련 성과물 의무 관리 	2

[그림 3] 서울시 공공빅데이터 활성화를 위한 주요 과제