

공공데이터 이용 활성화를 위한 정책에 관한 연구[†]

홍연웅¹

¹동양대학교 경영관광학부

접수 2014년 6월 10일, 수정 2014년 7월 2일, 게재확정 2014년 7월 7일

요약

최근 세계 주요국들은 데이터를 또 하나의 자원으로 인식하고 있으며, 정부의 투명성과 효율성을 제고하고 경제 활성화 효과를 기대하기 위하여 공공데이터를 적극 개방하는 정책을 경쟁적으로 시행하고 있다. 본 연구에서는 공공데이터의 개념을 알아보고 우리나라 공공데이터 정책의 현재와 문제점 및 공공데이터 이용을 활성화하기 위한 전략에 대하여 논의한다. 구체적으로는 데이터 품질의 중요성과 데이터 세트의 통일화 및 표준화 필요성, 공공 데이터 플랫폼 전략, 정부기관 간 또는 정부기관 내 공공 데이터의 공유 필요성 및 데이터 산업의 중심축으로서의 공공데이터와 빅데이터의 통합 지원 정책의 필요성 등이다.

주요용어: 공공데이터, 공공데이터 생태계, 플랫폼.

1. 서론

공공데이터를 민간에 개방하는 오픈데이터 정책이 세계적으로 확산되고 있다. 오픈데이터는 정부가 공공데이터를 개방하여 경제 활성화의 원동력으로 삼으려는 정책으로 2013년 6월 북아일랜드 G8 정상 회담에서 선진 8개국 정상들은 오픈데이터에 동참하겠다는 오픈데이터 헌장에 서명한 바 있다. 현재 세계 40여개 국가가 오픈데이터 포털 사이트를 운영하고 있으며, UN, EU, OECD 차원에서도 공공데이터 개방과 관련된 사이트를 개설할 예정이라고 한다. 미국이 오픈데이터정책을, 영국은 정보경제전략을, 일본이 전자정부 오픈데이터 전략 등의 정책을 시행하며 공공데이터 플랫폼을 구축하는 등 적극적인 개방 정책을 추진하고 있다. 공공 부문의 데이터 공개를 통해 행정의 효율성을 높이고, 국민의 참여를 활성화시키며 경제 활성화 등의 파급을 기대하고 있다. 정부의 데이터 공개 정책은 정보화 시대에 소통과 공유, 협업 전략이 무엇보다 중요하다는 것을 의미한다.

공공데이터 공개는 과거에도 추진되었지만 특히 최근에 주목받는 이유는 빅데이터 분석기술의 발달과 스마트 생활환경의 확산에 기인한다고 볼 수 있다. 과거에는 연구나 정책 개발의 목적으로 연구자나 업무 담당자가 공공데이터를 제한적이고 폐쇄적으로 활용했으나 최근에는 대중교통정보 서비스 등과 같이 많은 사람들이 수시로 자신의 정보단말기로 공공데이터를 활용하고 있다. 기업들도 신규 사업에 진출하거나 기존 사업의 경쟁력 제고 및 리스크 관리에 공공데이터를 유용한 자원으로 활용하고 있다. 18세기 까지 토지와 노동을 생산요소로 인식하다가 산업화가 진행되면서 자본이 더해졌듯이 정보화시대에는 전통적 생산요소에 데이터가 새로운 경제자원으로 추가되고 있다고 볼 수 있다. 이러한 배경에는 다양한 원시데이터 형태가 대량으로 광범위하게 제공되고 있으며 API (application programming interface), 링크드 데이터, CSV (comma separated value), 기계적으로 판독 가능한 (machine-readable) 형식 등

[†] 본 논문은 2014년도 동양대학교 연구비에 의하여 수행되었음.

¹ (720-711) 경북 영주시 풍기읍 교촌리 1, 동양대학교 경영관광학부, 교수. E-mail: hong@dyu.ac.kr

과 같이 사용 편의성이 향상되었을 뿐만 아니라 실시간으로 발생하는 데이터를 수집·분석·서비스 할 수 있는 정보통신기술을 꼽을 수 있다.

EU에서는 공공데이터 개방 정책의 기대효과로 활용 시장 규모로 500억 달러에 이를 것이라 추산하고 있다 (Vickery, 2011). 영국에서는 구체적으로 정부조직 운영 효율성 제고로 130~220억 파운드, 예산 절감, 세금오납, 조세포탈 감소로 10~30억 파운드, 세원 확대 등 추가 세수로 20~80억 파운드, 총계 160억~330억 파운드의 효과를 볼 것으로 추정하고 있다 (Yiu와 Kelsey, 2012). Heo 등 (2013)은 공공정보의 경제적 파급효과를 추정하고, 중요한 전·후방 연결 산업들을 파악하였다. 공공데이터의 비주얼라이제이션 (Lee, 2013)과 공공부문 빅데이터 활용을 주제로 한 연구 (Lee, 2013) 등이 있다. 이 외 공공데이터 민간개방 확대를 위한 법률 제정의 필요성을 주장하며, 공공데이터 민간 개방에 따른 경제 효과 분석을 시도한 Bai 등 (2013)의 연구가 있다. 정성적 데이터로부터 정량 데이터를 추출하고, 맥락을 쉽게 파악하고, 다양한 데이터세트를 연결하고 조합하는 방식에 대한 연구 (Faniel 등, 2012)와 데이터 생산자, 재사용자, 데이터 큐레이터와 데이터세트 자체 등 각 주체의 목적에 따라 데이터의 가치가 달라지며, 이러한 차이를 극복하기 위해 양질의 데이터세트 제공 기반을 마련해야 한다고 주장한 연구 (Faniel과 Yakel, 2011) 등이 있다.

본 연구에서는 실무적인 차원에서 국내 공공데이터의 이용 활성화를 위한 몇 가지 방안을 제시한다. 데이터 품질의 중요성과 데이터의 통일화 및 표준화, 공공 데이터 플랫폼 전략, 정부기관 간 또는 정부기관 내 공공 데이터의 공유 필요성 및 데이터 산업의 중심축으로서의 공공데이터와 빅데이터의 통합 지원 정책의 필요성 등에 대하여 연구하였다.

2. 공공데이터의 개념

공공데이터란 데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광 (光) 또는 전자적 방식으로 처리된 자료 또는 정보를 말한다 (공공데이터의 제공 및 이용활성화에 관한 법률 제2조 제2호). 따라서 공공데이터는 개별 공공기관이 일상적 업무수행의 결과물로 생성 또는 수집·취득한 다양한 형태 (텍스트, 수치, 이미지, 동영상, 오디오 등)의 모든 자료 또는 정보를 의미한다. 동법이 시행되기 전에는 각 개별 법령에서 정보, 기록물, 행정정보, 공공정보 등의 유사한 용어를 사용되어왔는데, 국가정보화기본법에서는 정보를 “특정 목적을 위하여 광 (光) 또는 전자적 방식으로 처리되어 부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식”이라고 정의했으며, 공공기관의 정보 공개에 관한 법률에서는 “공공기관이 직무상 작성 또는 취득하여 관리하고 있는 문서·도면·사진·필름·테이프·슬라이드 및 그밖에 이에 준하는 매체 등에 기록된 사항”이라고 정의했다. 공공기록물 관리에 관한 법률에서는 기록물을 “공공기관이 업무와 관련하여 생산하거나 접수한 문서·도서·대장·카드·도면·시청각물·전자문서 등 모든 형태의 기록정보 자료와 行政博物”이라고 정의했으며, 전자정부법에서는 행정정보를 “행정기관 등이 직무상 작성하거나 취득하여 관리하고 있는 자료로서 전자적 방식으로 처리되어 부호, 문자, 음성, 음향, 영상 등으로 표현된 것”이라고 정의했다.

공공데이터는 데이터의 공유와 재사용을 통해 가치가 실현된다고 할 수 있다. 공공 데이터란 활용, 재사용, 재배포가 누구에게나 허락된 데이터로 사용가능성과 접근성이 확보되어야 한다는 것이다. 데이터 전체의 사용이나 재생산에 추가 비용이 발생하지 않아야 하고, 편리하게 사용하고 수정할 수 있는 형식이어야 한다. 또한 재사용과 재배포가 가능해야 한다. 다양한 데이터세트의 혼합사용을 포함하여 데이터의 재사용과 재배포가 어떠한 제약도 없이 허용되고 참여자격 및 참여목적에도 제한을 두지 않고 누구나 보편적으로 참여하도록 개방적이다. 개인이나 단체에 상관없이 누구나 차별 없이 사용할 수 있어야 한다. EU는 2003년 11월 공공정보의 재사용에 관한 지침을 제정하고 각 회원국으로 하여금 적극

적으로 공공정보의 재사용에 관한 법률을 제정할 것을 권고하고 있다. 영국은 EU의 지침에 의거하여 2005년 공공정보 재사용에 관한 규정 (re-use of public sector information regulation)을 제정하고 공공정보의 재사용을 “공공기관이 보유하고 있는 공공정보가 해당 공공기관의 공무상 업무의 최초 목적과는 다른 목적으로 개인 또는 법인에 의해 사용되는 것”이라 정의한 바 있다. 또한 공공데이터는 데이터의 사양이 공개되어 자유롭게 이용 가능하고 기술적인 제약이 없는 형식으로 제공되어야 하며, 저작물에 부속되는 권리들은 재배포되는 모든 사람에게 모두 같이 적용되어야 하는데, 이 목적을 위해 어떠한 추가적인 라이선스도 요구하지 않는다. 특정 패키지에만 사용을 제한하는 라이선스를 부여하거나 다른 저작물의 배포를 제한하는 라이선스도 금지하고 있다.

3. 공공데이터 공개 현황

2014년 5월 24일 현재 공공데이터 포털 (www.data.go.kr)에 공개된 제공기관별, 분류체계별, 확장자별 공개건수를 보면 Table 3.1에서 Table 3.4와 같다. 제공기관별 공개건수는 총 8,028건 가운데 지방자치단체가 전체의 60%를 차지하고 있다. 입법기관, 교육기관 각종 위원회 등은 10건 이하로 미미한 실정이다. 분류체계별 공개건수는 일반공공행정이 전체의 27.1%로 가장 많고, 그 다음으로 산업 및 중소기업, 문화체육 관광, 사회복지의 순이다.

확장자별 공개건수를 보면 XLS와 XLSX가 48.3%로 가장 많고, 그 다음으로 HWP, HTML, JPG, PDF 파일 형태의 순서로 공공데이터라기 보다 공공문서로 추정되는 파일들도 적지 않음을 알 수 있다.

Table 3.1 Number of open data sets by public administrations

classification	number of data sets	classification	number of data sets
local governments	4,818	legislationorganizations	10
national administrative agencies	1,522	educational institutions	6
public institution	1,352	committees and free economic zone authorities	6
education administration agencies	314	total	8,028

Table 3.2 Number of open data sets by classifying system

classification	number of data sets	classification	number of data sets
general public administrations	2,183	regional development	382
industry/small and medium-sized businesses	769	public order and safety	305
culture, sports and tourism	726	agriculture, forestry and maritime	303
social welfare	687	science and technology	258
public health	638	telecommunication	143
education	628	national defense	72
environmental protection	443	national assembly	72
transportation and traffic	394	unification and foreign affairs	35
total			8,038

Table 3.3 Number of open data sets by file name extensions

file name extension	number of data sets	file name extension	number of data sets	file name extension	number of data sets
XLS	2,103	PDF	276	TEXT	27
HWP	1,560	TXT	212	WMV	21
XLSX	781	XML	173	JSP	19
HTML	326	CSV	115	ZIP	18
JPG	290	SHD	35	AVI	18
		total			5,974

공공데이터 이용실적 Table 3.4를 보면 최근 2년간 Open API 신청자는 4.7배, 활용건수는 10배 증가하였으나 대부분의 수요가 시내버스 도착정보 서비스에 집중되어 있어 수요가 다양하지 않음을 알 수 있다. 더욱이 기상정보제공 관련 기업을 제외하면 공공데이터를 이용한 창업 성공사례를 소수에 불과한

데, 이는 경제적 부가가치를 창출할 가능성이 높은 데이터가 거의 없거나 개방 초기에 시장형성 미흡 등으로 인해 창업 성장 환경이 취약한데서 그 원인을 찾을 수 있다.

Table 3.4 Using results of open government data portal site of Korea based on API demanders

classification	2011	2012	2013 (Jan. – Oct.)	total
number of demanders (persons)	1,446	2,629	6,755	10,830
Using results (million)	120	1,092	1,275	2,487

*Data source: www.data.go.kr

4. 이용 활성화 방안

본 장에서는 공공데이터 이용을 촉진시키기 위한 기술적 방안과 정책적 방안 및 조직 측면에서 고려할 사항을 제시한다.

4.1. 데이터 품질 관리

공공데이터 개방정책의 성패는 데이터 품질에 달려있다고 하여도 과언이 아니다. NIA (National Information Society Agency, 2011)의 조사에 따르면 공공 데이터의 정비 및 관리 필요성을 느끼고 있으나 제도적으로 근거조항이 없어 예산획득 혹은 내부에 사업 필요성에 대한 공감대 형성이 어려우며, 공공정보 민간 활용을 촉진하고자 정책을 펴도 보유하고 있는 데이터의 품질문제에 대한 자신감이 없어 공개할 수 없다는 의견이 다수를 차지하는 것으로 나타났다. 또한 공공기관 보유 DB의 정확성, 일관성, 유효성, 적시성 관점에서 데이터 품질진단을 한 결과 DB의 평균 오류율이 5.19%로 민간의 데이터 오류율 2.1% (NIA, 2011)를 상회하는 것으로 조사된바 있다. 이러한 품질문제를 해결하기 위해서는 첫째, 조직 목표에 부합하게 데이터 관리기준체계 마련과 품질관리정책, 역할과 책임을 명확히 한 추진체계의 마련이 필요하다. 둘째, 관리차원에서는 계획, 구축, 운영 등 데이터 수명주기 관점에서 데이터 품질을 통제할 수 있도록 품질 표준화, 품질진단 및 개선활동, 데이터 오류신고처리 등에 대한 데이터 관리프로세스를 마련하여 실무 담당자에게 충분히 교육되어야 한다. 그리고 품질관리를 위해 자동 품질진단 시스템 등 기술적인 인프라 구축도 동시에 검토되어야 할 것이다. 한편, 정부에서는 공공데이터 품질관리 지원 포털 (www.gooddata.kr)을 통하여 데이터 품질관리에 관한 종합지식제공, 각 기관의 품질관리체계 및 DB의 품질수준에 대한 온라인 자가진단, 전문기술상담 등을 실시하고 있어 향후 데이터 품질 개선에 중심 역할을 수행할 것으로 기대된다.

4.2. 데이터의 통일화 및 형식의 표준화

공공데이터 포털 (www.data.go.kr)에서 공개되는 대부분의 데이터는 데이터 생산자 또는 관련 업무 담당자의 판단에 의해서 결정되는 것이 대부분이다. 그러다보니 공개대상과 공개형식 등이 다양하고 통일화되어있지 않다. 지자체에 따라 동일한 내용이라도 형식이 상이하거나 공개여부에 대한 판단기준도 다른 경우도 있다. 결국 이용자의 불편을 초래하여 이용 활성화에 장애요인이 될 가능성이 높다. Table 3.1에서 보면 지자체의 데이터 공개건수가 가장 많지만 전국단위의 통합된 데이터라기보다 지자체별 서로 상이한 종류의 데이터로 생산한 지자체에만 적용할 수 있는 데이터가 대부분이기 때문에 사용 가치가 제한적이라고 할 수 있다. 또한 해설서, 업무 규정 등과 같이 기관 홈페이지의 게시판이나 공지사항으로 게시해도 될 문건을 데이터 포털에 공개하는 경우도 적지 않은 것으로 미루어보면 데이터의 사용가치보다는 공개건수에 관심이 적지 않음을 알 수 있다. 이러한 문제를 해소하려면 건수위주의 양적공개정책보다 부가가치중심의 공개방향으로 전환하고 공개대상을 체계적으로 결정하고 공개형식을 일괄되게 실시할 필요성이 제기된다.

4.3. 플랫폼 통일화를 통한 생태계 구축

세계적으로 많이 사용되는 대표적인 공공데이터 포털에는 3 가지 표준이 있다. 유럽형이라고 할 수 있는 CKAN과 미국형에 해당되는 Socrata 및 OGPL (open government platform)이다.

Socrata (www.socrata.com)는 미국의 연방정부 (www.data.gov)와 주정부들이 사용하는 민간 플랫폼으로 공공기관이 데이터를 저장, 관리, 공개, 활용하는 모든 프로세스를 제공한다. 미국 정부 기관들은 공통 플랫폼을 이용함으로써 데이터 공개 프로세스의 효율성을 제고하고 예산절감 및 정부기관 간에 데이터 연계를 실현하고 있다. 데이터세트 등록 과정은 데이터의 조사 및 식별, 데이터 정리 및 포맷 설정, 그리고 메타데이터 처리 및 등록의 3단계를 거치는데, 조사 및 식별 단계에서는 시민들에게 유용할 만한 데이터세트를 선별하게 된다. 국가 안보와 관련된 자료를 제외하고 전부 공개 대상으로 식별하는 것이 일반적이다. 정리 및 포맷 설정 단계에서는 데이터 베이스 테이블과 뷰 (view)를 스프레드시트로 처리할 수 있는 CSV나 웹에서 처리 가능한 XML 등의 다양한 형태로 추출하는 작업이 이루어지며, 이후 데이터 자체와 데이터세트의 제공 목적 등을 기술하게 된다. 이러한 세 단계 과정을 통해 원데이터를 등록하면 외부 이용자들에게 공개할 준비를 마치게 된다. Socrata 플랫폼에서 데이터세트의 분류는 각 기관의 분류체계에 따라 생산부서 및 주제분야로 나누어 설정되고, 활용성을 높이기 위한 방편으로 Socrata의 주제가 추가된다. Socrata는 다음에 설명한 CKAN과 OGPL에 비하여 가장 많은 기능을 제공하고 있으며, 시각화나 온라인 분석 툴, 매 쉬업 등 자동화 도구 기능을 갖추고 있다.

CKAN (www.ckan.org)은 공공데이터 개방 및 확산을 지원하는 오픈 소스 방식의 오픈 플랫폼으로, 비영리 단체인 Open Knowledge Foundation에 의해 개발되었다. 영국 (www.data.gov.uk)을 필두로 유럽 국가를 중심으로 약 40여개 국가에서 사용되어지고 있다. 한국의 www.data.go.kr도 CKAN 플랫폼을 사용하여 개발되었다.

CKAN은 국가나 지방자치단체 등의 공공 기관의 데이터 공급자들이 보유하고 있는 데이터를 공개하여 활용되도록 하는 것을 목표로 개발되었다. 데이터 수집이나 추가 기능 등이 필요할 경우 사용자별로 확장이 가능하다. 또한 데이터세트를 참조하는 응용 프로그램의 접근을 위해 풍부한 API를 제공한다. 데이터세트의 등록은 특정 포털 사이트가 CKAN 플랫폼을 사용하는 지방자치단체로부터 데이터세트를 획득하거나, 여러 CKAN 소스에 존재하는 특정 주제의 데이터세트만을 발췌하여 가져오는 것이다. CKAN은 데이터 색인 메타데이터로 DCAT (data catalog vocabulary) 표준을 사용하므로, CKAN 플랫폼을 사용하지 않는 색인으로부터 자료를 획득할 수 있다. 데이터세트를 공개하기 위해서는 웹 인터페이스를 통해 직접 데이터세트를 등록하거나 CKAN의 인터페이스나 자체 개발한 스프레드시트 추출 프로그램 등을 사용할 수 있다. 제공하는 메타데이터 요소는 제목, 고유 식별자, 기술요소, 갱신 내역, 미리보기, 라이선스, 태그, 제공 그룹, 포맷, API key, 기타 요소 등 11개 필드로 구성된다. CKAN은 콘텐츠 관리시스템과 연계가 가능하고 링크드 (linked) 데이터를 지원하여 확장성이 뛰어나다.

OGPL (www.ogpl.gov.in)은 미국과 인도정부가 공동으로 개발하여 2013년 1월에 공개한 플랫폼이다. 세계 각국의 공공데이터 개방 정책을 지원하기 위해 미국의 플랫폼인 www.data.gov의 오픈소스 형식의 베타버전이다. OGPL은 데이터의 등록 및 저장, 분석 및 시각화 기능 등 서비스에 꼭 필요한 기본 기능에 충실한 것을 알 수 있다.

이상의 대표적인 공공데이터 플랫폼을 통하여 정부 및 공공기관의 공공정보를 통합하여 서비스하고 있음을 알 수 있다. 국내에서도 특히 지방자치단체에서 공공데이터 서비스를 준비할 때 참고할 만한 시사점을 정리하면 다음과 같다.

첫째, 동일한 플랫폼의 사용이다. 공공 데이터세트의 가치를 높이려면 활용 가능한 데이터를 최대한 많은 시민들이 쉽게 사용할 수 있도록 해야 한다. 이를 위하여 각 지방자치단체나 공공기관이 구축할 플랫폼은 중앙정부 차원의 공공데이터 포털과 기능을 연계할 수 있도록 해야 할 것이다. 둘째, 데이터세트

가 실제 서비스되기 위한 다양한 솔루션이 요구된다. 일반 사용자를 위한 온라인 시각화 도구와 서비스 담당자를 위한 Open API 변환도구가 대표적이다. 셋째, 데이터 서비스 결과는 지방자치의 훌륭한 정책 도구가 될 수 있다는 측면에서 지역 특성에 부합한 데이터를 발굴하는 것이 필요하다. 서울시의 시내 버스도착 서비스 정보나 심야버스노선 개발이 좋은 사례이다. 시민의 정보 요구를 면밀히 분석하여 그에 부합하는 지역 밀착형 공공데이터를 개발하는 것이 중요하다. 넷째, 공공데이터 포털에 대한 접근성 제고이다. 미국과 영국의 공공데이터 포털은 별도의 로그인 절차 없이 공공정보에 접근해 데이터를 내려 받을 수 있다. 한국의 공공데이터 포털에서 데이터를 내려 받으려면 로그인 절차를 거치고 필요데이터를 신청해야 한다. 시스템 안정성과 보안성을 강화한 대신 접근성과 활용성을 희생시켰다고 할 수 있다.

정부는 데이터를 공개하고 서비스 기업과 시민은 데이터를 사용하여 새로운 가치를 창출한다. 이러한 과정에서 참여 주체사이에 가치 있는 교류와 피드백이 형성되고, 정부의 투명성과 효율성이 증가하게 된다. 공공데이터 플랫폼은 열린 정부 데이터를 내외부와 상호연결해 주는 플랫폼이며, 데이터 포털이라는 핵심 기능을 수행하고 있다. 이러한 공공데이터 부문에서 서비스로서의 플랫폼 (PaaS; platform as a service)이라는 것을 통하여 정부가 제공하는 데이터 시장에서 공공정보 제공자와 수요자가 모여 다양한 방법으로 활용되는 Figure 4.1의 공공데이터 생태계를 구축하는 것이다.

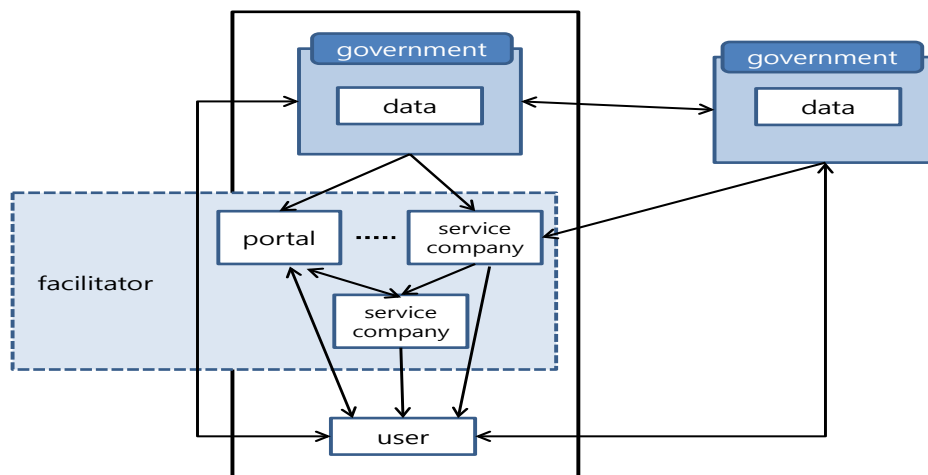


Figure 4.1 Ecosystem of open government data

4.4. 적극적 데이터 공개 정책 추진

공공데이터 개방정책의 핵심가치는 정부의 투명성 확보와 시민의 참여제고 및 경제적 부가가치 창출에 있다. 이를 달성하기 위해서는 사회적 수요가 높은 데이터를 적극 발굴하고 공개하려는 전향적인 정책이 요구된다. 정부기관이 보유중인 데이터의 공유정책이 절실하다. 예를 들어 국세청, 국토교통부, 보건복지부가 보유하고 있는 개인의 재산이나 과제, 복지 및 건강보험관련 정보를 상호 공유할 경우 신규 세원 발굴이나 세금의 과·오납 문제해결, 복지사각지대 발굴 및 보조금의 부정 수령 예방 등에 활용될 수 있어 정부의 효율성과 사회정의 실현에 기여할 수 있을 것이다. 한 기관 내부에서도 서로 다른 부서가 보유하고 있는 내부 데이터를 공유할 경우에도 이와 유사한 효과를 기대할 수 있을 것이다. 제한된 범위 내에서 데이터를 공유하는 정부기관 내 또는 정부기관 간 보유 데이터를 공유하고 연계 협력을 촉진할 수 있는 제도적 장치를 적극 검토할 필요성이 요구된다.

4.5. 탑다운식 운영 조직 구축

공공데이터 개방 정책은 안전행정부의 정부 3.0정책의 일환으로 탑다운 (top-down) 방식으로 추진되고 있다. 특정 데이터의 경우는 탑다운식으로 수요조사에 의해서 개방 대상이 결정되고 해당되는 하부기관이 관련 데이터를 공개하는 경우도 있다. 적지 않은 데이터의 경우 개방대상과 개방형식의 결정이 일선의 담당자에 의해 이루어지는데, 수요조사에 의한 판단보다는 보유하고 있는 데이터의 완성도나 업무의 편의성 위주로 이루어지고 있다. 즉 정책 수립은 탑다운 방식, 공개 대상 데이터의 결정은 실무자에 의한 바텀업 (bottom-up) 방식으로 이루어지므로 개방 데이터세트의 수량에 비하여 사용가치는 이에 미치지 못하고 있는 실정이다. 데이터의 사용가치를 제고하기 위해서는 탑다운 방식과 바텀업 방식의 혼합하는 현행 개방 방식을 탑다운 방식으로 전환·운영할 필요가 있다.

4.6. 빅데이터와 공공데이터 연계 지원 정책

빅데이터 관련 정책은 연구개발 기능을 가진 미래창조과학부가 수행하고, 공공데이터 관련 정책은 연구개발 기능이 없는 안전행 정부가 관장하고 있어 데이터 산업의 육성 관점에서 공공데이터 정책이 본연의 기능을 발휘하기 어려운 구조이다. 즉, 연구개발 예산의 지원이 제한적이거나 거의 없기 때문에 공공데이터 이용 활성화 정책에 한계가 있다고 볼 수 있다. 그러므로 민간부문이 필요에 의하여 자발적으로 이용하지 않으면 공공데이터를 적극적으로 활용할 유인 요인이 거의 없는 실정이다. 빅데이터와 공공데이터를 연계하여 활용할 수 있는 정책을 적극 발굴하여 추진할 경우 정책의 시너지 효과가 확대될 것이다. 구체적 대안으로는 공공데이터의 개방은 안전행 정부가 관장하고 이용 등 산업적 부문은 미래창조과학부 또는 산업통상자원부가 담당하는 것이 바람직하다고 사료된다. 공공데이터도 빅데이터 이상으로 데이터 산업적 가치를 가지고 있음을 인식하고 적극적인 지원정책을 추진해야 할 것이다. Kim과 Cho (2013)는 빅데이터의 응용분야와 빅데이터 자료 분석에 유용하게 사용되는 통계적 방법론들에 대하여 소개하였는데, 이러한 방법론들은 공공데이터 분석에도 동일하게 활용된다.

5. 결 론

각국 정부는 공공정보를 개방하여 새로운 경제적·사회적 부가가치를 창출하고 있다. 중앙정부 및 지방정부와 공공기관들은 공공데이터 플랫폼을 사용하여 공공 데이터의 효율적인 수집 및 확산을 도모하고 있다. 우리나라도 안전행정부의 공공데이터포털이 공공자원 확산의 포털 역할을 하고 있다. 하지만 공개된 데이터의 경제적 부가가치 창출이 미흡한 실정이고 이용의 불편함 등으로 서비스의 효율성은 개선될 여지가 적지 않다.

이러한 상황에서 본 연구는 공공데이터의 이용을 활성화하기 위한 몇 가지 방안을 제시하였다. 먼저, 공공데이터 품질관리의 중요성과 이에 대한 지속적 관리활동을 강조하였다. 지자체별로 공개대상 데이터세트를 통일화하고 공개 형식을 표준화할 것을 제안하고 양 위주의 공개정책보다 부가가치중심의 공개 정책으로 전환할 것을 주장하였다. 외국의 대표적 공공 데이터 플랫폼인 Socrata와 CKAN, OGPL을 통해 서비스 플랫폼의 역할과 기능을 살펴보고, 우리나라의 공공 데이터 서비스 현황과 비교하여 시사점을 분석해 보았다. 정부기관 내 각 하부부서 또는 정부기관 간 보유 데이터를 상호 유기적으로 공유하고 연계 협력을 촉진 할 수 있는 제도적 장치를 적극 검토할 것을 강조하였으며, 공공데이터 관련 조직을 전면적으로 탑다운 방식으로 운영할 것을 제안하였다. 끝으로 안전행정부와 미래창조과학부로 분리되어있는 공공데이터와 빅데이터 업무를 데이터 활용 부문만 통합 운영할 필요성을 제기하였다. 본 연구의 의의는 빅데이터와 더불어 데이터 산업의 양대 축인 공공데이터의 이용활성화를 위해 이루어져야 할 실무적 내용을 제시한데 있다. 정부기관 내 또는 정부기관 간 보유 데이터를 공유할 법·제도적 대안을 구체적으로 제시하지 못한 점이 본 연구의 한계라 할 수 있다.

References

- Bai, S. H., Lee, J. Y., Song, S. H., Jang, J. B., Kang, S. K., Yoon, J. S., Lee, D. H. and Kim, J. H. (2013). A study on the legislation necessity for activation of public data. *Journal of Korean Association for Regional Information Society*, **16**, 67-86.
- Lee, M. J. (2013). Public data analytics and visualization. *Journal of Korean Institute of Information Scientists & Engineers*, **31**, 101-107.
- Lee, S. H. (2013). A big data example to support public services. *Proceedings of Korean Institute of Information Technology Summer Conference*, Korean Institute of Information Technology, 121-123.
- Faniel, I. and Yakel, E. (2011). Significant properties as contextual metadata. *Journal of Library Metadata*, **11**, 155-165.
- Faniel, I., Kriesberg, A. and Yakel, E. (2012). Data reuse and sensemaking among novice social scientists. *Proceedings of the American Society for Information Science and Technology*, **49**, 1-10.
- Heo, P. S., Park, K. M., Park, W. J., Cho, K. S. and Ryu, W. (2013). Domestic market and economic impact of the re-use of public sector information. *Electronics and Telecommunications Trends*, **28**, 118-131.
- Kim, Y. and Cho, K. H. (2013). Big data and statistics. *Journal of the Korean Data & Information Science Society*, **24**, 959-974.
- Ministry of Government Legislation. (2013). *The law on provision and use of public data*, available from <http://www.law.go.kr/main.html>.
- Ministry of Government Legislation. (2014). *The law on information disclosure of a public institution*, available from <http://www.law.go.kr/main.html>.
- National Information Society Agency. (2011). *Information strategy report for public information quality improvement*, National Information Society Agency, Seoul.
- Vickery, G. (2011). *Review of recent studies on public sector information re-use and related market developments*, European Commission Directorate- General for Information Society, EC.
- Yiu, C. and Kelsey, T. (2012). *A right to data-fulfilling the promise of open public data in the UK*, Policy Exchange, UK.

A study on the invigorating strategies for open government data[†]

Yeon Woong Hong¹

¹School of Management and Tourism, Dongyang University

Received 10 June 2014, revised 2 July 2014, accepted 7 July 2014

Abstract

Recently many countries have established open government data platforms to disclose government or government controlled entities-owned data that can be freely used, reused and redistributed by anyone. Open government data can help you to make better decisions in your own life, or enable you to be more active in society. Open data is also making government more effective and transparent, which ultimately also reduces costs. This paper explains the open data concepts and circumstances in Korea, and also suggests detailed invigorating strategies such as data quality policy, data unification and standardization policy, open data service platform, and integrated support plan of big data and open government data.

Keywords: Ecosystem of open data, open government data, platform.

[†] This research was supported by Dongyang university research fund.

¹ Professor, School of Management and Tourism, Dongyang University, Kyungbook 720-711, Korea,
E-mail: hong@dyu.ac.kr