

우리나라 공공데이터의 이용활성화 방안에 관한 연구: 링크드 오픈 데이터화 전략을 중심으로

A Study on Revitalizing the Use of Korean Public Data: Focused on Linked Open Data Strategy

이현정 (Hyun Jung Yi)*

남영준 (Young Joon Nam)**

초 록

우리나라는 공공데이터 제공과 관련된 제도가 최근 제정되면서 정부기관 및 지방자치단체 등의 공공기관이 보유한 데이터를 적극적으로 개방하고 제공하는 방향으로 정책이 변화하고 있다. 개방의 목적은 크게 두 가지로 구분한다. 정부운영의 투명성을 확보하여 국민의 알 권리를 충족시키는 것이다. 다른 하나는 공공데이터를 하나의 국가부존 자산으로 활용하여 국익을 창출하기 위함이다. 이 연구에서는 공공데이터의 개방 현황을 분석하고 개선방안을 제시하였다. 연구범위는 지방자치단체에서 제공하는 공공데이터이기 때문에 서울특별시를 비롯한 17개 광역시도와 기초 자치단체 228개 시·군·구에서 보유한 것을 전수 조사하였다. 연구결과에 따르면 지방자치단체는 각 기관에서 생산 및 소장한 공공데이터에 대한 목록과약과 공개에 대해 상대적으로 소극적인 것과 개방 데이터의 포맷도 특정 소프트웨어에 의존적인 형태였다. 이러한 점을 해결하기 위해서는 궁극적으로 지역 공공데이터개방과 활용을 높일 방안으로 링크드 오픈 데이터 형태로 개방하는 필요성과 방안을 제시하였으며, 국가 공공데이터개방을 위한 통합 플랫폼을 통한 종합적 개방절차와 방안을 제안하였다.

ABSTRACT

In South Korea, systems related to the provision of public data were recently implemented. As a result, policy changes have been made that are headed in the direction of actively providing open access to data held by public institutions, such as government agencies and local municipalities. The purpose of the open I will be divided into two broad. To ensure the transparency of government operations, and is intended to satisfy the right to know the people. The other one is to create national interest by utilizing the public data as one country endowment assets. In this study, we analyze the open situation of public data, were presented the improvement measures. Range of research, the public data that local government owns, to determine to have a central information and other limitations and characteristics, Seoul the beginning to the seventeen regional support municipality 228 that you have held for city districts were census. According to the research results, local governments, themselves produced, is a relatively reluctant to disclosure and understanding of the list of public data that are holdings. According to the research results, local governments, themselves produced, is a relatively reluctant to disclosure and understanding of the list of public data that are holdings, also emphasizes the conservative value than take advantage of value have had. Therefore, it was determined that there is a need to resolve several issues through disclosure via a linked data format as a strategy to increase the openness and utilization of local public data.

키워드: 지역공공데이터, 링크드 오픈 데이터, 정부정책
local public data, linked open data, government policy

* 중앙대학교 문헌정보학과 박사수료(caulis98@gmail.com) (제1저자)

** 중앙대학교 문헌정보학과 교수(namjy@cau.ac.kr) (교신저자)

- 논문접수일자: 2014년 11월 20일 ■ 최초심사일자: 2014년 12월 1일 ■ 게재확정일자: 2014년 12월 22일
- 정보관리학회지, 31(4), 249-266, 2014. [http://dx.doi.org/10.3743/KOSIM.2014.31.4.249]

1. 서론

세계 각국은 정부운영과정의 산출물로서 생산되는 공공데이터를 새로운 미래 국가자원으로써 가치를 중시하고 있다. 일반적으로 각국 정부들이 판단하는 공공데이터의 가치는 두 가지로 대별할 수 있다. 하나는 공공데이터를 통해 정부에서 수행하는 모든 업무를 국민에게 공개함으로써 정부운영의 투명성을 확보하는 가치이다. 다른 하나는 국민이 개방된 공공데이터를 활용하여 국민과 기업들이 공공데이터 부가가치를 극대화하여 이익을 창출할 수 있는 재화적 가치이다. 따라서 주요 국가들은 공공데이터의 개방을 통해 정부운영의 투명성 확보와 산업적 일자리 창출에 국가의 역량을 집중하고 있다.

우리나라도 예외는 아니어서 박근혜 정부는 '정부 3.0'이라는 국정 아젠다를 제시하고 정부가 보유하고 공개할 수 있는 모든 공공데이터를 원칙적으로 전부 개방하는 정책을 집행하고 있다. 이를 위해 '공공데이터 이용활성화에 관한 법'과 같은 법적 기반을 마련하고, 행정자치부에서 '공공데이터포털'과 같은 시스템적 기반도 마련하였다. 우리나라는 이와 같은 법적 및 시스템적 지원에 따라 공공데이터 포털을 통해 2014년 9월 현재 10,031건의 공공데이터셋을 개방하여 여러 분야에서 이를 활용하고 있다.

그 개방의 수준은 정부업무에서 생산되는 워드파일이나 엑셀파일과 같은 문서작성의 결과물을 공개하는 수준과 함께 데이터의 연결을 기계로 처리할 수 있는 수준까지 다양한 형태로 개방하고 있다. 투명성을 확인하기 위한 공공문서의 원래 형태(original shape)로 표현하고 있는 문서파일과 산업적 활용을 높이기 위해 공공

문서를 트리플 구조와 같은 링크드 오픈 데이터(LOD) 형태로 개방하는 것이다. 링크드 오픈 데이터는 공공데이터에 HTTP와 RDF, URIs 등과 같은 국제표준의 주소정보와 의미정보를 추가로 부착하여 데이터 간 결합성을 극대화한 데이터 형태이다. 정부가 공공데이터를 링크드 오픈 데이터 형태로 개방하는 이유는 거대한 공공데이터를 언어나 소장 기관에 상관없이 웹을 통해 누구나 접근하여 재활용할 수 있는 여건을 제공하기 위함이다. 이러한 장점 때문에 영국을 비롯한 많은 선진 국가들은 링크드 오픈 데이터의 형태로 공공데이터를 발행하여 그 활용성을 극대화하는 정책을 시행하고 있다.

이에 비해 우리나라의 공공데이터 개방건수는 중앙정부에서 1,689건(전체 17%)과 지방자치단체에서 5,923건(59%)에 달하여 개방의 수준은 선진국 수준이다. 그렇지만 이 가운데 링크드 오픈 데이터 형태로 개방한 공공데이터는 없기 때문에 링크드 오픈 데이터와 같이 공공데이터 개발에 따른 활용성을 높이기 위한 전략과 방안이 절대적으로 부족한 상황이다.

이 연구는 이 점에 착안하여 우리나라 지방자치단체에서 링크드 오픈 데이터를 구축하기 위해 표준화된 변환방법과 기준을 공공데이터 발행단계에 적용하기 위해 고려할 요인을 파악하고자 한다. 이 연구에서는 현재 지방자치단체에서 개방하고 있는 데이터셋을 전수조사하여 각 데이터셋의 개방 포맷을 분석한다. 또한 전문가 인터뷰를 통해 지역공공데이터의 링크드 오픈 데이터 발행을 위한 방안을 도출한다. 궁극적으로 전수조사결과와 인터뷰 결과를 바탕으로 링크드 오픈 데이터 발행을 위한 방안과 전략을 제안한다.

2. 이론적 배경

2013년 10월부터 「공공데이터의 제공 및 이용 활성화에 관한 법률」이 공공정보 개방과 공유 확대를 위해 시행되고 있다. 이 법에서 “공공기관”은 국가기관, 지방자치단체 및 「국가정보화기본법」 제3조 제10호에 따른 공공기관을 의미한다. “공공데이터”는 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. “제공”이란 공공기관이 이용자로 하여금 기계 판독이 가능한 형태의 공공데이터에 접근할 수 있게 하거나 이를 다양한 방식으로 전달하는 것을 말한다. 이와 같이 법에서 명시한 공공데이터를 개방하는 플랫폼은 공공데이터포털을 비롯하여 서울열린데이터광장, 경기도공공데이터공개포털, 지방행정데이터 개방 등이 있다. 또한 공공데이터베이스(데이터 세트)는 시도행정정보시스템을 비롯한 서울행정시스템(구, 시군구행정정보시스템), 행복e음 시스템과 같은 업무용 행정시스템 등을 통해 생성되고 있다.

2.1 공공데이터플랫폼

2.1.1 공공데이터포털

공공데이터포털¹⁾은 공공기관이 생성 또는 취득하여 관리하고 있는 공공데이터를 한 곳에서 제공하기 위한 통합창구로서 행정자치부에서 관리하고 있다. 2010년 6월 국가지식포털 내에 공

공정보 개방 및 활용과 관련된 공공과 민간의 다양한 개방요구를 지원하기 위한 목적으로 공공정보활용지원센터가 설치되었고 공공정보목록 안내서비스와 공공정보 신청안내 등의 서비스를 시작하였다. 한국정보화진흥원은 공공데이터포털과 공공정보활용지원센터를 위탁운영하고 있다. 2014년 9월 14일 현재 공공데이터포털은 704개의 국가기관, 지방자치단체, 기타 공공기관, 민간을 통한 10,031개의 데이터세트를 개방하는 공유서비스로 성장하였다. 이 가운데 중앙정부에서 공공데이터의 개방건수는 1,689건(전체 17%)이며 지방자치단체에서 공개한 지역공공데이터는 5,923건(59%)이다. 공공데이터포털은 데이터 활용을 위한 단계와 조건, 데이터 제공을 위한 절차 및 방법, 기술표준에 대하여 다양한 방법으로 안내하고 선도하는 역할을 수행하고 있다.

2.1.2 서울열린데이터광장

서울시는 서울시에서 파생되는 공공데이터를 적극적으로 공개하는 데이터관리 정책을 적용하고 있다. 이는 서울시민에게 서울시 운영의 투명성을 확보하기 위한 방안으로 열린 시정 2.0이라는 캐치프레이즈로 나타나고 있다. 서울시 공공데이터 기반의 민간중심 서비스 창출과 함께 새로운 형태의 공공가치 창출을 목표로 설정한 정책이다. 서울시 공공데이터를 정보 공개적 가치와 함께 활용적 가치를 제고하기 위해 서울열린데이터광장²⁾이라고 명명된 자체 공공데이터 개방 플랫폼을 구축하였다. 서울열린데이터광장은 서울시민의 생활과 밀접한 교통, 문화 등

1) 공공데이터포털. Retrieved from <http://data.go.kr>

2) 서울열린데이터광장. Retrieved from <http://data.seoul.go.kr>

과 관련한 공공데이터를 공개한 일련의 개방자장공간이다. 이 플랫폼을 통해 서울시는 2014년 9월 현재 1,438건의 공공데이터셋을 생산 및 보유하고 있다. 개방된 공공데이터의 유형은 Sheet를 비롯하여 Chart, Map, Open API, File, Link 등 다양한 형태의 공개수준으로 구성되어 있다. 한편 서울열린데이터광장에서 적용한 공공데이터 분류체계는 국내외 여러 공공데이터 분류체계를 분석하여 대분류로써 교통을 비롯하여 안전, 환경, 일반행정, 도시관리, 보건, 복지, 산업경제, 문화관광, 교육의 10개 분야로 구분하고 있다. 이를 근거로 분야별로 총 29개의 하위분류를 두고 있다. 또한 서울시는 공공데이터 개방지침을 비공개 대상을 제외하고 모두 개방함을 원칙으로 하고, 시민이 편리하게 이용할 수 있도록 표준화되고 다양한 데이터 제공방식을 통해 제공하는 것, 공공데이터의 신뢰도 향상을 위해 데이터 품질을 강화하는 등의 10가지 개방원칙(서울열린데이터광장, 2014)을 천명하고 있다.

2.1.3 경기도공공데이터공개포털

경기도는 경기도공공데이터공개포털³⁾을 통해 경기도에서 생산되는 데이터셋을 공개하고 있다. 이 포털은 경기도에서 파생되는 공공데이터에 대한 개방 요청을 받으면 이를 제공하는 형식으로 운영하고 있다. 시스템적으로 경기도에서 생산하는 공공데이터를 망라적으로 전면 개방하는 정책과 함께 창구로 요청되는 의견을 통해 실시간 또는 주기적으로 제공 가능한 공공데이터를 발굴하여 제공하는 개방정책을 사

용하고 있다. 단순파일 요구나 1회성 자료에 대한 요청은 '정보공개청구(<http://info.gg.go.kr>)'를 이용하도록 하여 공공데이터 개방과 정보공개청구를 정확하게 구분하여 서비스를 제공하고 있다. 경기도공공데이터공개포털을 통해 제공하는 분류체계는 조세/법무/행정을 비롯한 9개로 공공데이터를 구분하고 있다. 플랫폼은 영국의 CKAN 플랫폼을 기본으로 이용하고 있으며, 현재 서비스의 수준은 BETA로 설명하고 있다. 2014년 9월 현재 '경기도내 통통신장 현황 데이터'를 비롯한 77건의 데이터셋이 CSV 파일 형태로 등록되어 있다. 또한 데이터개방 목록을 통해 '경기도 자동차 등록현황'을 비롯한 522개 파일을 공개하고 있으며, 파일의 형태는 한글과 PDF, 엑셀 등의 형태로 구성되어 있다. 해당 파일들은 특정한 로그인의 과정 없이 누구나 해당 포털에 접속한 이용자들이 획득할 수 있다.

2.1.4 지방행정데이터개방

지방행정데이터개방⁴⁾은 행정자치부(중앙정부)의 주관으로 지방자치단체가 보유·관리하고 있는 공공데이터 가운데 국민생활과 밀접한 데이터를 개방하는 데이터 공개플랫폼이다. 이 플랫폼의 특징은 우리나라 지방자치단체에서 보유하고 있는 공공데이터의 개방을 집중적으로 취급하는 곳으로써, 향후 의료, 보건 등 48분야 530여 종의 공공데이터를 개방할 예정이다. 2014년 9월 현재 지방행정데이터개방에서는 일반음식점, 관광업소 정보에 대한 2종의 데이터셋을 시범적으로 개방하고 있다. 향후 개방대

3) 경기도공공데이터공개포털. Retrieved from <http://data.gg.go.kr>

4) 지방행정데이터개방. Retrieved from <http://www.gmap.go.kr/tcportal/saeol/SaeolInfo.do>

상시기와 개방목록의 수는 새올행정시스템을 비롯한 지방자치단체에서 관리하고 생산하는 공공데이터 전체로 확대할 예정이다. 특히 해당 플랫폼에는 등재된 공공정보의 재가공을 통한 새로운 서비스 모델의 창출부터 자유로운 이용을 저작권법과 같은 법적인 구속력 없이 자유롭게 활용한다는 규정을 명시하고 있다. 이는 사업을 통한 이익창출에 정부차원의 적극적 지원을 명시한 것이다. 이 시스템은 현재 2개의 데이터세트만을 공개하고 있으며, 해당 데이터 세트에 탑재되어 있는 목록은 목록형식으로 제공하지 않고 있다. 이는 실시간성 데이터 제공보다는 해당 데이터에 대한 주기적인 업로드에 의해 이루어질 수 있기 때문에 사용자의 편의성이나 개방의 규모는 다른 개방 플랫폼에 비해 상대적으로 작은 수준이다. 또한 공개 시계열에 대한 원파일(raw file)을 망라적으로 개방하는 방식이 아니라 시계열을 지정하고 해당 기간에 공공데이터의 존재를 검색결과로 예측하는 형태로 이루어져 있다. 실질적인 공공데이터 개방 플랫폼보다는 향후 지방자치단체에서 생산하는 공공데이터를 개방하는 방법을 효율적으로 도출하기 위한 시범적 플랫폼의 성격이 강하다.

2.2 지방행정시스템

2.2.1 행정정보시스템

우리나라의 행정정보시스템은 전자정부 구현의 일환으로 국가와 지방자치단체 차원에서 추진되었다. 국가차원에서는 입법부, 사법부, 행정부의 중앙부처 및 산하단체에서의 행정사무를 수행하기 위해 추진되었으며 지방자치단체

는 개별 시도와 시군구 및 산하단체에서 행정사무를 처리하기 위해 추진되었다. 국가차원에서 지방자치단체에 위임해야 할 업무가 상당히 많은 중앙집권적인 우리나라의 행정체계상 업무의 효율과 개별 지방자치단체의 시스템 중복개발을 방지하기 위해 구현된 시스템이라 할 수 있다(강두호, 2011). 행정자치부에서 개발·운영하고 있으며, 17개 광역 시도 단위에서 사용하는 시도행정정보시스템과 228개 기초자치단체에서 사용하는 새올행정시스템으로 나뉜다. 행정정보시스템은 국가위임사무를 처리하고, 지방자치단체의 행정사무를 표준화하고, 대 주민 민원처리 업무를 효율적으로 처리하는 것을 주요한 목적으로 개발·보급되었다. 시도행정정보시스템이나 새올행정시스템에서 개방하고 있는 데이터는 지방행정데이터개방에서 2종에 대한 데이터세트를 서비스하고 있고, 공공데이터포털에 등록된 데이터세트도 있지만 해당 현황을 명확히 파악할 수 없는 상태이다. 2014년 10월 현재, 시도행정정보시스템에서 97종, 새올행정시스템에서 358종의 지역공공데이터가 개방될 예정으로 데이터의 목록이 공개되어 있다.

2.2.2 행복e음시스템

행복e음시스템(사회보장정보시스템)은 기존의 행정자치부 새올행정시스템에서 복지분야를 분리하여 각종 복지 급여 및 서비스 지원 대상자의 자격 및 이력에 관한 정보를 통합·관리하여 지방자치단체 복지업무 처리를 지원하기 위해 구축된 시스템이다. 행복e음시스템은 표준화된 복지사업 프로세스에서 양산되는 복지정보, 각 복지사업별 정보, 상담사례관리 정보로 구분

할 수 있으며 이를 근간으로 복지정보통계시스템을 구축하여 행복e음시스템을 통해 통계서비스를 제공하고 있다. 2014년 10월 현재, 104종의 지역공공데이터가 개방될 예정으로 데이터 목록을 공개하였다.

2.3 링크드 오픈 데이터 발행의 장점

미국 등을 포함하여 각 나가는 공공데이터를 수집하여 이를 링크드 오픈 데이터의 형태로 변환하여 제공하고 있다. 이러한 과정에서 축적된 지식 및 개발·사용한 도구들을 출판하고 공개하였다(Ding, 2011; Ding, 2010). 이들은 기존의 공공데이터를 링크드 데이터로 변환하는 데 있어서 자주 겪는 어려움으로 (1) 서로 다른 데이터세트에서 각각 다른 메타데이터의 구조를 사용하거나 통일되지 않은 용어를 사용하는 문제, (2) 데이터가 정제되지 않아 서로 다른 데이터를 자동화의 방법으로 병합하는 데 한계가 있는 점 등을 지적하였다. 링크드 오픈 데이터는 데이터를 출판하기 위한 일련의 원칙으로서 데이터를 서로 연결할 수 있도록 하여 데이터를 더욱 쉽게 활용할 수 있다. Archer 등의 보고서를 통해 링크드 오픈 데이터로 공공데이터를 서비스할 때의 장점을 다음과 같이 정리할 수 있다.

첫째, 신속성 및 적절성: 정보를 재사용하고자 하는 소비자가 그들이 필요로 한 때에, 그들이 필요로 하는 데이터에 접근할 수 있도록 하는 환경을 제공해 준다. 따라서 더 이상 전체 데이터세트나 데이터베이스를 내려 받아 사용할 필요가 없이 온라인상에서 필요한 때에, 필요한 데이터에 접근하여 사용할 수 있다.

둘째, 비용대비 효과 상승: 서로 다른 데이터 공급자들은 자신들의 데이터를 링크드 데이터 서비스의 기반구조를 통해 동일한 링크드 데이터의 형태로 데이터를 서비스할 수 있어 상당한 경제적인 이익을 줄 수 있으며 비용을 절감할 수 있다. 한편, 데이터 이용하는 소비자 입장에서는, 서로 다른 데이터를 병합하여 재사용하기 위해서 더 이상 로컬의 저장소에 필요한 데이터세트 전부를 저장할 필요가 없으므로 데이터의 저장과 관리에 드는 비용을 절감할 수 있다. 뿐만 아니라, 서로 다른 데이터세트로부터 온 데이터를 병합하여 사용하기 위한 과정 중, 각 데이터의 매핑의 작업을 컴퓨터에 의해 자동화할 수 있으므로 기존에 이 과정에 소요되던 인적, 시간적 자원을 절감할 수 있다.

셋째, 데이터 질의 향상: 링크드 데이터는 데이터의 질과 관련하여 다음과 같은 긍정적인 효과를 가져 온다. (1) 데이터가 지속적이며 안정적인 URIs를 통해 제공되므로 데이터 내의 객체를 참조하는 데 있어서 의미의 모호성을 해결할 수 있다. 또한 (2) URIs에 기반한 식별자를 동일한 객체를 가리키는 다른 식별자와 연결하여 사용할 수 있어 그 의미를 보다 분명히 하여 다른 데이터와 쉽게 병합하여 데이터의 의미를 풍부하게 만들 수 있다.

넷째, 데이터의 최신성을 유지할 수 있다. LOGD (Linked Open Government Data)를 이용하는 데이터 재사용자는 특정 데이터세트에 대해 하나의 버전만이 존재하게 되므로 언제든지 가장 최신 버전의 데이터에 접근하여 이를 이용할 수 있다.

3. 지역공공데이터의 개방현황

본 연구에서 지역공공데이터는 서울특별시, 부산광역시, 대구광역시, 광주광역시, 인천광역시, 대전광역시, 울산광역시, 세종특별자치시, 경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도, 제주도의 17개 시도 지방자치단체와 산하 228개 시·군·구 기초 자치단체에서 생산·개방하여 제공하는 데이터로 정의한다. 2014년 9월 14일 기준으로 공개된 공공데이터포털 5,923건, 서울열린데이터광장 1,438건, 경기도공공데이터포털 77건

의 총 7,438건의 지역공공데이터를 전수 조사하였다. 전체 개방현황을 제공형태별로 분석하였다.

3.1 지역공공데이터 개방 포맷

현재 정부에서 개방하여 제공하고 있는 공공데이터의 포맷 종류는 원시파일의 형태에서부터 RDF 형태에 이르기까지 매우 다양하게 분포되어 있는데, 공공데이터포털, 서울열린데이터광장, 경기도공공데이터공개포털에서 제공하고 있는 데이터형태를 <표 1>과 같이 구분하였다.

<표 1> 데이터 제공 형태별 특징

형태	특징
그리드	Socrata의 'Dataset' 유형, 서울열린데이터광장에서는 Sheet(서울시는 공공데이터포털에 비해 훨씬 많음)에 해당하는 것으로 원천 데이터셋을 웹브라우저에서 스프레드시트 형태로 출력해주는 방식, CSV, JSON 등의 파일로 변환하여 저장할 수 있음
차트	웹페이지 안에서 막대그래프, 선 그래프 등의 도표로 표현해주는 데이터. 사용자가 막대 형태로 접이나 선 형태 등을 반영하여 원하는 형태의 그래프 만들 수 있음
지도	위치정보 데이터셋을 지도와 매쉬업 하여 보여줌. 위치데이터의 수가 많을 경우 화면을 출력하는 시간이 오래 걸릴 수 있음
다운로드	가장 많은 비중을 차지하는 유형으로 텍스트(txt, doc, hwp), 이미지(jpg, gif, bmp) 등 다양한 형태의 파일형식으로 돼 있고 확장자명을 기재함. 서울열린데이터광장에서는 'File'이라는 유형으로 구분된 형태로서 동일하게 파일을 다운로드받는 방식이다. 이러한 유형은 특정 소프트웨어에 의존적인 데이터셋이 대부분이다.
Open API	다운로드 다음으로 큰 비중을 차지하는 유형으로서 많은 개발 자원이 소요되지만 응용 프로그램을 개발하는 등 민간 활용도가 높은 서비스이다. 일정한 형식으로 공개해서 응용 프로그램을 개발할 수 있도록 함
LINK	LINK는 데이터셋이 플랫폼에 등록되지 않고 데이터셋이 실제 존재하는 해당하는 기관 사이트로 연결하는 링크가 제공하는 것을 의미한다.
개방예정	데이터셋에 대한 정보만 확인할 수 없고 파일이 등록되어 있지 않거나 내용은 현재 확인 불가능
배송 및 기타	등록 불필요(기관 배송)하며, 관리자에게 통보되어 공공데이터를 요청한 사람에게 직접 배송하는 것을 의미함
LOD	링크드 오픈 데이터는 데이터를 연결하여 웹을 확장할 수 있는 가장 실제적인 방법이다. 링크드 오픈 데이터는 자유로운 재사용은 제한하지 않는 오픈 라이선스 하에서 발행된 링크드 오픈 데이터이다. LOD인 경우 해당 발행 사이트의 URL을 포털에 등록하여 연결함
RDF	RDF 구문을 따라 구조화하여 데이터를 표현하기 위해서 흔히 XML을 이용하여 RDF의 형식으로 변환하거나 N-triples와 같이 보다 단순한 트리플 구조로 데이터내의 객체간의 관계를 정의하는 방법으로 데이터를 제공함

총 10가지의 형태로 공공데이터를 개방하고 있는데, 공공데이터포털, 서울열린데이터광장, 경기도공공데이터공개포털의 구분기준이 약간 상이하였으나 공공데이터포털을 기준으로 유형에 맞게 통합하여 분류하였다. 이 플랫폼을 통해 제공하는 개방형 파일 포맷은 하나의 데이터세트에서 2개 이상의 형태로 제공하는 것을 권장하고 있으며 목적에 맞게 자유로운 활용을 담보할 수 있는 사용권리에 대해서도 명시적으로 선언해야 한다.

3.2 지역공공데이터 개방 현황

전체 7,438건의 지역공공데이터는 <표 2>와 같은 포맷으로 개방되어 있다.

<표 2> 지역공공데이터 개방현황

개방 포맷	개방 건수	비율(%)
그리드	863	9.545
차트	69	0.763
지도	227	2.510
다운로드	4,113	45.492
Open API	878	9.711
LINK	2,714	30.018
개방예정	14	0.154
배송 및 기타	53	0.586
LOD	4	0.044
RDF	106	1.172
합계	9,041	100

<표 2>의 합계와 전체 데이터세트 수가 다른 것은 2개 유형 이상의 포맷으로 제공되는 데이터를 해당 유형에 각각 합하였기 때문이다. 9,041건 중 1,603건, 약 18% 정도가 2개 이상의

개방형 포맷으로 공개하고 있다. 전체적으로 다운로드 형태가 약 45%로 가장 큰 비중을 차지하고 있고, LINK, Open API, 그리드 순으로 조사되었다. 링크드 오픈 데이터 형태는 총 4건으로 조사되었으나 실제 링크드 오픈 데이터 형태로 명기된 데이터세트는 링크드 오픈 데이터 형태로 발행되지 않았다. 모든 데이터들은 실제 해당 사이트를 조사하면 관련 링크드 오픈 데이터로 표현한 데이터세트를 설명하는 필드의 텍스트(description)로 인해 공공데이터포털에 잘못 분류된 경우였다. 또한 공개한 공공데이터의 유형을 분석하면 다운로드를 비롯하여 LINK, Open API와 같은 형태로 개방되어 있다. 이 경우에 LINK는 향후 링크드 오픈 데이터로의 변환이 사실상 불가하며 링크드 오픈 데이터로의 연결도 활용면에서 매우 부적절한 형태 가운데 하나이다. 또한 다운로드의 경우도 해당 데이터에 대한 축적의 수준이기 때문에 활용성이 상대적으로 떨어진다. 즉 다운로드 형태로 제공하는 파일 가운데 PDF나 그림 파일 등과 같은 형태의 파일은 활용을 위해 적합하지 않기 때문에 발행의 수준을 적극적으로 안내하여야 한다. 예를 들면, PDF나 그림 파일로의 개방은 지방행정 운영에 대한 공개 이상의 의미를 갖지 못하기 때문에 활용적 의미의 개방을 위해서는 최소한 RDF로 변환이 가능한 수준의 포맷 형태로 개선하는 정책이 필요하다.

4. 링크드 오픈 데이터 발행 전략

공공데이터 개방에 관한 담당자의 인식을 전

환하고 제공된 데이터셋을 능동적으로 활용할 수 있는 환경을 조성하고자 전문가 자문을 통해 구체적인 방안을 모색하고자 하였다.

4.1 전문가 자문

공공데이터전략위원회의 발표에 따르면 공공데이터 활용을 저해하는 요인으로 저작권 문제를 비롯하여 법과 제도 기반의 미비를 가장 중요한 것으로 조사되었다(공공데이터전략위원회, 2014). 대표적으로 대부분의 공공기관 담당자들은 공공정보를 제공함으로써 업무처리의 실수나 미비를 노출시켜 불필요한 불이익을 받을 수 있다는 부담 때문에 공공데이터 개방에 대해 소극적인 형태를 갖는다. 따라서 이러한 문제점을 해결하기 위해서는 공공데이터 활용적 측면과 함께 공무원의 적극적 개방, 운영시스템에 대한 해결 방안 등과 같은 제반 문제에 대한 전문가 의견이 필요하다. 왜냐하면 문제의 해결은 제공자로서 공무원이나 혹은 사용자로서 국민들의 일방적 의견만으로는 접근점을 찾기 어렵기 때문이다. 따라서 이러한 점을 고려하여 이 연구에서는 2014년 10월 13일부터 17일까지 전문가 3명을 대상으로 공공데이터 관한 전문가 심층면접 조사를 실시하였다. 해당 전문가는 국내에서 링크드 오픈 데이터를 발행한 경험이 있는 업체 담당자와 정부 3.0의 공공데이터 개방을 총괄하고 있는 부서의 담당자를 대상으로 하였다. 심층면접 조사의 방법으로는 공공데이터 개방과 앞으로의 방향에 대한 자문을 구하는 형식을 취하였으며, <표 3>과 같이 대상, 형태, 방법에 대해 현재의 공공데이터 개

방현황과 앞으로 방향에 대한 의견, 어떤 형태로 개방해야 할 것인지에 대한 발행형태, 플랫폼에 관한 의견으로 나누어서 정리하였다. 이를 통해 향후 지역공공데이터 발행 전략에 대한 전문가 의견을 구하였다.

이상에서 도출한 지역공공데이터 활용상의 문제점을 정리하면 다음과 같다.

첫째, 사용자가 동일한 주제의 데이터셋을 활용하고자 할 때 개방하고 있는 지방자치단체와 개방하지 않고 있는 지방자치단체가 존재하므로 국가차원의 전국적인 데이터 확보가 필요한 경우에 완전한 데이터 확보가 어렵다.

둘째, 지방자치단체별 개방의지와 인식도가 달라서 공개가 가능한 데이터셋인데도 불구하고 해당 데이터셋의 존재를 확인할 수가 없는 경우가 발생하고 있다. 데이터셋에 있어서 보존적 가치를 중요시한다면 목적에 따른 데이터셋의 범위와 종류가 구분되어야 할 것이다.

셋째, 국가행정기관과 지방자치단체의 업무가 다르므로 지역공공데이터는 국가행정기관에서 보유한 공공데이터와는 다른 특성을 가지고 있을 것이다. 지역공공데이터 내에서 발달한 다른 주제의 데이터셋에 대한 발굴이 필요하다.

넷째, 공공데이터 개방 포맷 중 원시파일에 해당하는 데이터들이 대부분을 차지하고 있었다. 전문가 자문을 통해 많이 지적되었던 바처럼, 공공데이터의 활발한 이용을 위하여 이기종간의 연계가 가능한 포맷으로 발행되어야 하므로 공공기관이 링크드 오픈 데이터로 발행하는 것을 적극적으로 유도할 필요가 있다.

〈표 3〉 전문가 자문면접 결과

	구분 (LOD: 링크드 오픈 데이터)	내용
전문가 1	대상: 개방현황 및 목록	<ul style="list-style-type: none"> • 공공데이터포털에 공개되고 있는 목록은 동일한 내용에 대해서 모든 기관에서 일괄 제공되지 않아서 새로운 서비스를 만들려고 해도 데이터셋이 완전하지 못한 어려움이 있다.
	형태: 제공포맷 - LOD로 발행할 데이터셋의 우선순위, LOD로 발행해야 하는 이유	<ul style="list-style-type: none"> • 우리나라는 5 star의 2레벨 수준이 가장 많은 비율을 차지하고 있는 것으로 알고 있다. 하지만 활용을 위해서는 레벨 4, 5 이상으로 끌어올려야 한다. • 그러기 위해서는 여러 가지 교육을 시행하고 시사점을 많이 주어야 유인이 될 수 있을 것이다.
	방법: 플랫폼	<ul style="list-style-type: none"> • 기존 데이터베이스 및 API를 활용하여 어려움 없이 공개할 수 있는 기반시스템의 개발과 보급이 필요하다. • 보유 데이터를 변환하여 단계별 검증을 거친 데이터를 선별하여 공개하는 것도 방법이 될 수 있다.
전문가 2	대상: 개방현황 및 목록	<ul style="list-style-type: none"> • 일반 사용자는 현재 개방되고 있는 정도가 정부데이터의 어느 정도를 개방하고 있는 것인지, 개방되지 않은 목록은 어떤 것들이 있는지에 대한 내용을 전혀 알 수 없다. • 개방되지 않은 목록 중에서 새로운 서비스를 개발할 만한 가치가 지닌 데이터셋이 있다고 하더라도 이에 대해서 인지할 수 있는 기회가 현재로서는 없다.
	형태: 제공포맷 - LOD로 발행할 데이터셋의 우선순위, LOD로 발행해야 하는 이유	<ul style="list-style-type: none"> • 설문조사 및 공청회 등 일반 사용자들이 참여할 수 있는 기회의 제공이 중요하다. • Open API로 개발된 데이터 및 서비스를 우선적으로 LOD로 발행하여 최종 사용자의 의견을 피드백을 받아 개선해 나가야 할 것이다.
	방법: 플랫폼	<ul style="list-style-type: none"> • 공공데이터포털 내의 지역공공데이터가 서울시나 경기도 등 지방자치단체가 각각 갖고 있는 데이터셋과의 연계가 부족하다. • 공공데이터포털 내에서는 확인할 수 없으나 플랫폼을 통해 개방되어 있다는 사실을 확인할 수 없다.
전문가 3	대상: 개방현황 및 목록	<ul style="list-style-type: none"> • 데이터 보유 기관들이 각자 데이터를 독립적인 예산으로 플랫폼을 개발하여 발행하여 표준을 잘 지키고 있는지 검증할 수 있는 장치나 방법이 없다. • 우리나라는 현재 API에 비하여 LOD 개방이 현저히 낮은 비율로 발행되어 있고, 시범적으로 발행하는 기관도 생겨나고 있다. 해외 사례에 비추어 볼 때 우리나라의 LOD 개방현황은 걸음마 단계를 벗어나지 못한 수준이므로 데이터 보유 기관들의 LOD 발행을 적극적으로 유도할 필요가 있다.
	형태: 제공포맷 - LOD로 발행할 데이터셋의 우선순위, LOD로 발행해야 하는 이유	<ul style="list-style-type: none"> • 이용자가 즉시 활용 가능한 실생활에 적극 활용될 수 있는 데이터의 개방이 우선되어야 하고, 제공형태 또한 스파클 등 어려운 문법을 가진 RDF 뿐만 아니라 XML, JSON 등 이기종간의 쉬운 연계가 가능한 포맷으로 발행할 수 있어야 효율적인 이용을 권장할 수 있을 것이다.
	방법: 플랫폼	<ul style="list-style-type: none"> • Virtuoso 등 세계적으로 많이 쓰이고 있는 발행 플랫폼을 사용하여 DBPedia 등과의 효율적인 연계를 위해 발행플랫폼 자체 개발을 지양해야 한다.

4.2 지역공공데이터 발행 전략

4.2.1 공공데이터 목록의 공개

공공데이터포털에서 원하는 공공데이터를 제

공하지 않는 경우 '공공데이터 제공제도'에 의거하여 공공데이터 제공을 신청할 수 있다. 그러나 공공데이터의 존재유무를 파악할 수 없는 상태에서 이 제도를 적극적으로 이용할 수 있도록

활성화하는 것은 어렵다.

개방되지 않은 목록 중에서 새로운 서비스를 개발할 만한 가치가 지닌 데이터셋이 있다고 하더라도 현재로서는 국민들이 이를 인식할 수가 없다. 시도행정정보시스템과 서울행정시스템을 통해 생산되는 공공데이터 중 현재 공개되고 있는 것을 중심으로 공공데이터포털, 서울열린데이터광장, 경기도공공데이터포털, 지방행정데이터개방 등의 서비스가 이루어지고 있다. 그러나 개방되고 있는 것 이외의 현재는 지방행정시스템에 탑재되어 있는 전체 지역공공데이터의 수량을 파악할 수 없는 상태이다. 행정정보시스템에 탑재되어 있으나 비공개 대상인 것과 공개대상이지만 아직 데이터플랫폼에 탑재되지 않은 데이터를 비롯하여 공개할 수 있는 데이터셋을 발굴하고자 하는 노력이 필요하다. 개인정보보호법에 정의하는 민감정보를 암호화하거나 삭제하는 작업 등을 통해서 기술적으로 가능한 범위를 설정할 수 있다. 다만, 저작권법 24조 2(공공저작물의 자유이용)에 따라 국가나 지방자치단체가 공표한 저작물을 별도의 허락없이 이용할 수 있으므로, 공공누리의 청록색 마크가 부착된 데이터를 자유롭게 이용할 수 있다. 지방자치단체가 보유하고 있는 데이터 현황을 열람할 수 있도록 한다면 공공데이터에 대한 요구도를 향상시킬 수 있을 것이다. 제공 요청이 발생했을 때, 공개가 가능한 데이터인지에 대해 내부적으로 검토하여 공개불가로 판정할 수도 있을 것이다. 어떤 데이터에 대한 요구도가 있는지를 의견 수렴 절차를 거쳐서 우선적으로 링크드 오픈 데이터로 발행하거나 단순 개방할 목록의 양과 순서를 검토하는 작업을 통해서 진행할 수 있을 것이다.

4.2.2 공개 2수준의 개방

공공데이터의 개방은 국가전략으로 이루어지는 것이기 때문에 이에 따른 예산과 인력이 필요하다. 특히 링크드 오픈 데이터형태로의 개방은 관련 공공데이터 세트를 RDF화하는 것이 일반적이기 때문에 이에 대한 예산을 추가적으로 확보하여야 하는 절차가 필요하다. 따라서 많은 링크드 오픈 데이터 선진국에서도 개방하는 모든 공공데이터를 링크드 오픈 데이터 수준으로 개방하지 않으면서도 개방에 따른 효과는 높이며 개방에 따른 법적 문제는 없는 공공데이터 중심으로 LOD화하는 정책을 수행하고 있다. 한편 각국은 공공데이터 개방수준을 링크드 오픈 데이터 수준으로 개방하는 것을 권고하지만 실제적으로 개방의 수준을 5단계로 구분하여 개방의 수준보다 개방의 범위를 확대하는 전략을 수립하고 있다. 즉 가능한 많은 공공데이터를 개방하는 것을 지향하고 데이터의 특성과 활용도를 고려하여 LOD 수준으로 개방을 권장하고 있다. 우리나라는 전자정부환경의 도입으로 정부운영의 데이터들은 모두 전자형태의 데이터로 구축되어 있어 기본적으로 다음 <표 4>의 2단계 수준으로 이미 축적되어 있다.

<표 4> 5 Stars of LOD

수준	파일형태	비고
1 수준 ★	이미지	
2 수준 ★★	한글, 엑셀	
3 수준 ★★★	CSV	
4 수준 ★★★★	RDF	N-Triple
5 수준 ★★★★★	LOD	타LOD연계

위 <표 4>와 같이 공공데이터의 개방은 Open Data 공개 5 수준(흔히 5 stars)(한국정보화진흥

원, 2014; Bauer & Kaltenböck, 2012; Janowicz 등, 2014)으로 개방하는 것은 링크드 오픈 데이터 형태로 개방하는 것을 의미하고, 이와 같은 개방수준은 공공데이터의 활용성을 최대한 높일 수 있는 방식이다. 하나이 수준으로 고도화하는 것은 일련의 데이터정련과정을 거쳐야 하기 때문에 타 개방데이터와 링크드한 형태로의 5 수준 데이터는 상대적으로 많은 예산과 노력이 선행되어야 한다. 따라서 지방자치단체에서 시도행정정보시스템과 새울행정시스템에서 생성한 공공데이터의 개방을 5수준으로 하는 것은 현실적으로 한계를 갖는다. 따라서 우선적으로 현재 행정운영시스템을 통해 생성된 전자화된 데이터를 최소한의 가공 과정없이 생성된 형태(Born Publish)로의 개방을 수행하는 것을 우선적으로 고려해야 한다. 즉 일차적으로 개방을 2수준의 개방포맷형태로 개방하고, 데이터의 산업적 활용도를 고려하여 궁극적으로 타 데이터와 연계된 5수준으로 확대하는 전략이 필요하다.

4.2.3 Open API에서 LOD로 변환

공공데이터 데이터세트의 공개 유형은 일회성 데이터 개방 형태와 지속적 개방형태로 구분할 수 있다. 일회성 데이터세트는 특정한 업무영역에서 특정 기간 내 생성이 완료되는 경우이다. 예를 들면 2002 월드컵대회와 같이 특정 사안에 대해 정부의 경기장 건설관련 공공데이터세트들은 개최 전 기획 단계부터 개최 후 사업완료 보고에 이르기까지 일련의 시리즈(綴) 형태로 이루어진 데이터세트이다. 다른 하나는 정부운영에 반복적으로 파생되어 생산되는 데이터세트이다. 현재 지방자치단체에서

생산되는 데이터세트들은 대부분 후자의 경우와 같은 형태로 개방되고 있다. 따라서 지방행정정보시스템에 탑재되어 있는 공공정보들은 실시간으로 등록되는 정보이기 때문에 관리인력과 사용자 편의를 위해 API 방식으로 제공하는 것이 가장 효율적이다. 즉 제공자 위주의 공공정보제공 방식이기 때문에 어플리케이션의 주도권(governance)과 하위 이용자에 대한 서비스 영향력을 정부가 유지할 수 있다. 이용자 관점에서도 미리 구현된 데이터구조를 알고 있기 때문에 개발이 용이한 어플리케이션을 확보할 수 있다. 반면, 상기의 방법은 제공자 중심이기 때문에 실제 사용자가 원하는 데이터를 확보하기 어렵다. 이는 지방행정정보시스템의 모든 공개데이터를 DB구조에 있는 원 모델로 제공할 경우에는 문제가 발생하지 않지만 정부입장에서 필요한 정보만을 제공할 경우에는 사용자 요구를 파악하지 않으면 공개에 따른 활용성이 저하될 수 있다. 기존의 API로 제공하는 형태를 일차적으로 가공하여 유사자원과의 링크드 형태로 파일들을 배치로 처리하여 사용자에게 제공할 수 있다. Open API는 다양한 프로토콜 및 아키텍처 형태로 제공되지만 공공데이터개방을 위해서는 REST 기반의 Open API일 때 요청은 HTTP 메소드와 URL을 이용하고 응답은 XML과 같은 데이터 포맷을 기반으로 표현되기 때문에 입력과 출력에 대해서 공통된 규격으로 정의하는 정책이 필요하다(행정안전부, 2011). 공공데이터포털에서는 Open API 서비스 활용을 위해 공유서비스 활용자 클라이언트가이드를 제공하고 있다(행정안전부, 2014). Open API 및 가이드를 활용하기 위해서는 프로그래밍 지식이 필수이며, 서비스

코드에 가이드 코드를 적용하기 위해서는 여러 변수를 수정해야만 한다. 따라서 이용 활성화를 위해서는 위 가이드의 암호·복호화 및 코드를 언어별, 라이브러리 및 공통 모듈로 만들어 제공해야 공공정보 제공 및 활용을 촉진할 수 있을 것이다.

단, Open API는 서비스를 활용하기 위해 서비스 신청 후 키를 발급받고, 서비스 요청시 매번 키를 암호화하여 요청하고 응답받을 때는 복호화 하는 단점을 갖고 있다. 또한, 하나의 Open API에 접근해 서비스를 활용할 수 있지만 두 개의 데이터를 한 번에 접근해서 서비스를 만들거나 원하는 조건을 통해 해당되는 데이터에만 접근해서 서비스를 활용하기 어렵다. 두 개 이상의 데이터에 접근하여 데이터를 활용하려면 링크드 오픈 데이터로 가능한 단점도 갖고 있다. 그럼에도 불구하고 API를 통하여 데이터베이스로부터 직접 데이터를 가져올 수 있고 API에 의해 미리 정의된 특정 포맷의 데이터로 실시간으로 변환이 가능하기 때문에 여섯 단계의 절차⁵⁾를 거쳐야 하는 변환 사례보다 링크드 오픈 데이터로 발행하는 과정을 간소화할 수 있다. 즉 API를 이용하여 XML과 JSON 등과 같은 공개포맷으로 변환하면서 해당 데이터 객체를 내·외부의 URI로 정의하는 방법으로 쉽게 변환할 수 있도록 한다. 따라서 데이터를 단순히 개방 및 공유하는 것뿐만 아니라 개방된 데이터의 연계 활용을 위해 링크드 오픈 데이터로 발행하는 것이 필요하다. 또한, 링크드 오픈 데이터 변환비용으로 메타데이터 스키마 설명과 RDF화, URI부착 비용과 같은 추가 예산의 필

요성을 반드시 고려해야 한다.

4.2.4 통합 플랫폼의 활용

지역공공데이터는 행정자치부에서 운영하고 있는 공공데이터포털과 서울시, 경기도 각각에서 구축한 플랫폼과 더불어 음식점, 관광업소 정보만을 우선 개방하고 있는 지방행정데이터 개방으로 분산되어 있었다. 공공데이터포털에서 서울열린데이터광장에서 등재된 공공데이터파일은 일부 중복적으로 접근할 수 있기 때문에 사용자 관점에서는 하나의 플랫폼으로 접근하였으나 필요한 복수의 공공데이터를 확보하기 위해 복수의 플랫폼을 네비게이션해야 하는 문제점을 갖고 있다. 이와 같이 지방자치단체에서 자체적으로 구축한 데이터 개방 플랫폼을 통한 공공데이터제공은 이와 같이 사용자 관점에서는 불편함으로 작용할 수 있다. 따라서 현재 시도행정정보시스템을 통해 97종의 지역공공데이터와 서울행정시스템을 통해 358종의 지역공공데이터, 행복e음시스템을 통해 104종의 지역공공데이터가 개방할 예정이며 이는 '지방행정데이터개방'을 통해 관련 데이터들이 일괄 개방될 예정이다. 사용자 관점에서 최선의 방법은 하나의 플랫폼에서 원하는 공공데이터를 일괄적으로 입수하는 것이다. 이러한 방법은 지방자치단체의 역할과 같은 문제 때문에 실제로 공공데이터의 개방은 하나의 정부에서 다양한 채널로 이루어지고 있다. 한편 전문가 자문에서 밝혀진 바처럼, 지방자치단체의 개별 공공데이터 플랫폼을 제작하는 것을 지양하고 통합 플랫폼을 지향하는 것이 필요하다. 이러한 것은 지방자치단체와 중앙

5) 링크드 오픈 데이터 발행절차는 일반적으로 식별(identify), 모델링(model), 명명(name), 설명(describe), 변환(convert), 발행(publish)의 여섯 단계의 절차를 거친다.

정부간의 협력적 차원에서 중앙정부가 관련 공공데이터를 통합적으로 관리할 수 있는 메타공공데이터 관리 방안을 마련해야 한다. 한편 이러한 중앙정부의 노력과 함께 지방자치단체들도 공공데이터의 관리와 개방을 위한 자체적인 플랫폼으로써 CKAN 등과 같은 세계적으로 표준화된 플랫폼을 원용하는 전략이 필요하다.

4.2.5 분류체계의 통일성 유지

공공데이터의 활용을 높이는 것은 활용이 용이하도록 사용자 중심의 직관적인 인터페이스를 구축하는 것이 필요하다. 해외의 공공데이터 개방 플랫폼은 디자인적인 요소도 가능한 유사하게 설계하고 특히 분류체계의 일관성을 유사하게 설계하고 있다. 현재 우리나라의 정부와 지방자치단체에서 운영하는 플랫폼은 서로 다른 인터페이스를 사용하고 있다. 공공데

이터포털은 교육을 비롯한 16개의 분류체계로 초기화면에서 제공하고 있으며, 지방행정데이터개방은 관광업소와 일반음식점으로 2개의 분류체계를 사용하고 있다. 서울시의 경우는 교육을 비롯한 10개의 분류체계로 이루어져 있으며 대분류의 주제영역도 많은 차이를 보이고 있다. 이에 비해 경기도는 9개의 분류체계를 갖고 있으며, '소방/재난/안전'은 공공데이터포털과 주제영역의 크기가 차이를 보이고 있다. 또한 2014년 9월 현재 '농림/해양/수산'과 '교육/취업'에 관한 데이터세트는 한 건도 수록되어 있지 않다(<표 5> 참조).

공공데이터 포털과 서울시의 경우와 같이 메뉴의 체계와 주제영역이 동일한 혹은 유사한 개념어가 채택되면 이용자는 각 플랫폼에 대한 등록 데이터세트에 대한 예측을 통한 공공데이터 입수가 수월할 것이다. 또한 이와 같은 사용자

<표 5> 우리나라 주요 공공데이터 개방 플랫폼의 주제영역

플랫폼	공공데이터포털	서울시	경기도	지방행정데이터
주제 대분류	교육	교육	교육/취업	-
	국토관리	도시관리	도시/주택	-
	공공행정	일반행정	조세/법무/행정	-
	재정금융	-	-	-
	산업고용	산업경제	산업/경제	-
	사회복지	복지	-	-
	식품건강	-	-	-
	문화관광	문화관광	관광/문화/체육	관광업소, 일반음식점
	보건의료	보건	가족/보건/복지	-
	재난안전	안전	소방/재난/안전	-
	교통물류	교통	교통/환경	-
	환경기상	환경	-	-
	과학기술	-	-	-
	농축수산	-	농림/해양/수산	-
	통일외교안보	-	-	-
법률	-	-	-	

측면과 달리 공공데이터를 등록하는 업무를 맡은 공무원의 경우에도 분류체계 항의 직관성은 해당 공공데이터세트 등록의 정확성을 담보할 수 있기 때문에 오분류 가능성을 크게 감소시킬 수 있다. 이에 비해 공공데이터포털의 환경기상과 서울시의 환경은 유사하나 경기도의 경우는 교통영역에 환경을 대입하여 이의 불일치는 사용자의 혼란을 초래할 수 있다. 즉, 공공데이터 개방 플랫폼 분류체계의 통일성 유지와 정합성 유지를 통한 조정이 필요하다. 분류체계의 통일성 유지는 제공자 측면의 등록과 사용자 측면의 검색에 편의성 확보를 위해 가장 중요한 요소이다. 따라서 국가차원의 분류체계에 대한 가이드를 지방자치단체와 그 철학을 공유하는 것이 필요하다. 이를 통해 사용자가 필요한 공공데이터를 서로 다른 플랫폼에서 입수하는 경우에도 효율적인 공공데이터 입수를 가능하게 할 것이다.

5. 결론

정부나 혹은 공공기관이 생산하는 공공데이터는 크게 정보공개와 의미와 데이터 개방의 의미를 갖는다. 전자는 해당 기관의 투명성을 담보하는 것이고 후자는 산업적 활용성을 지원하는 것이다. 이 연구는 공공데이터의 산업적 가치를 제고하는 방안가운데 링크드 오픈 데이터화 전략을 연구하였다. 이를 위해 우리나라 공공데이터포털을 분석하여 공공데이터의 발행주체를 분석하였다. 분석 결과에 따르면 특히 우리나라는 공공데이터의 생산현황을 분석한 결과 지방자치단체가 중앙정부에 비해 2014년 현

재 약 3.6배 이상의 개방실적을 갖고 있었다. 이런 현상을 근거로 지방자치단체에서 생산하는 공공데이터를 공개하는 플랫폼으로서 공공데이터 포털과 서울열린데이터광장, 경기도공공데이터공개포털, 지방행정데이터개방의 특성을 분석하였다. 분석에 따르면 공공데이터포털은 704개의 국가기관을 비롯하여 지방자치단체, 기타 공공기관, 민간을 통한 10,031개의 데이터세트를 개방하고 있었다. 서울시는 1,438건의 공공데이터세트를 비롯하여 경기도는 77건의 데이터세트와 522개 파일을, 지방행정데이터개방은 2개의 데이터세트를 개방하고 있다. 또한 경기도와 지방행정데이터개방은 2014년 9월 현재 BETA 버전으로 플랫폼을 운영하고 있었다. 각 플랫폼의 개방데이터의 유형은 다운로드가 가장 많았으며 LOD의 형태는 가장 적은 것으로 조사되었다. 따라서 전문가의 자문을 통해 공공데이터의 LOD형태 발행을 저해하는 요인을 다음과 같이 도출할 수 있었다. 첫째, 지방자치단체별로 개방수준과 목록, 유형이 다르기 때문에 전국적인 공공데이터세트의 확보가 어렵다. 둘째, 지방자치단체별로 부존재정보가 존재하며 또한 개방의지와 인식도도 달라 이를 체계적으로 입수할 수 있는 채널확보가 어렵다. 셋째, 중앙정부의 데이터 제공 채널(공공데이터포털)과 지방자치단체의 특화된 지역공공데이터를 지속적으로 발굴할 수 있는 제도가 필요하다. 넷째, 공공데이터의 개방 포맷을 적극적으로 LOD화하는 것이 필요하다.

이상의 조사와 분석을 통해 우리나라 공공데이터 개방의 범위를 확대하고 해당 데이터의 산업적 가치를 높일 수 있는 방안을 다음과 같이 제안하였다.

첫째, 공공데이터의 목록공개. 일부 지방자치단체는 자체적으로 개방의 범위를 정하지만 또한 국민수요자의 요구를 받아 개방하는 정책을 갖고 있다. 사용자들이 어떤 데이터가 존재하는지를 파악함으로써 개방의 요구와 수요측정이 가능하다.

둘째, 공개 2수준의 개방 정책. 일차적으로 개방을 2수준의 개방포맷형태로 개방하고, 데이터의 산업적 활용도를 고려하여 5수준으로 확대하는 전략이 필요하다.

셋째, Open API에서 LOD로 변환정책의 필요성. 공공데이터를 개방단계부터 데이터의 연계 활용을 위해 링크드 오픈 데이터로 발행하는 전략과 절차가 필요하다.

넷째, 통합 플랫폼의 활용. 중앙정부의 노력과 함께 지방자치단체들도 공공데이터의 관리와 개

방을 위한 자체적인 플랫폼으로써 CKAN 등과 같은 세계적으로 표준화된 플랫폼을 원용하는 전략을 수립한다.

다섯째, 분류체계의 통일성 유지. 분류체계의 통일성 유지는 제공자 측면의 등록과 사용자 측면의 검색에 편의성 확보를 위해 가장 중요한 요소이다. 따라서 국가차원의 분류체계에 대한 가이드를 지방자치단체와 그 철학을 공유하는 것이 필요하다.

이상의 연구와 제안을 통해 중앙 및 지방자치단체에서 개방하는 공공데이터를 적극적으로 활용할 수 있는 기반을 조성할 수 있다. 궁극적으로 개방의 수준을 링크드 오픈 데이터로 발행하여 산업적 활용을 위한 중요한 국부자원화하는 장기적 전략수립방안연구도 필요할 것이다.

참 고 문 헌

- 강두호 (2011). 국가보급 행정정보시스템 서비스 품질이 지방자치단체 사용자 만족도에 미치는 영향. 석사학위논문. 부산 부경대학교 대학원.
- 경기도공공데이터포털. Retrieved from <http://data.gg.go.kr/>
- 공공데이터포털. Retrieved from <http://www.data.go.kr/>
- 서울열린데이터광장. Retrieved from <http://data.seoul.go.kr/>
- 안전행정부 (2014). 데이터로 만드는 새로운 세상. 공공데이터의 제공 및 이용 활성화에 관한 법률 시행 1주년 세미나. 공공데이터전략위원회.
- 정완 (2013). REST 기반 Open API의 입출력 파라미터 특성을 고려한 자동 매쉬업 방법에 대한 연구. 석사학위논문. 서울 광운대학교 대학원.
- 지방행정데이터개방. Retrieved from <http://www.gmap.go.kr/tcportal/saeol/SaeolInfo.do/>
- 한국정보화진흥원 (2014). LOD 기반의 데이터 관리 패러다임 전환 전략. IT & Future Strategy 보고서. 한국정보화진흥원.

- 행정안전부. 공유서비스 개발 가이드라인 (2011). Retrieved from http://www.mogaha.go.kr/cmm/fms/FileDown.do?atchFileId=FILE_00000000023299&fileSn=1/
- 행정안전부 (2014). 국가정보자원 개방·공유체계 구축 공유서비스 활용자 클라이언트 가이드.
- Archer, P., Dekkers, M., & Goedertier, S. (2013). Business models for linked government data
Retrieved from http://ec.europa.eu/isa/documents/study-on-business-models-open-government_en.pdf/
- Bauer, F., Kaltenböck, M. (2012). Linked open data: The essentials.
Retrieved from <http://www.semantic-web.at/LOD-TheEssentials.pdf/>
- Ding, L., DiFranzo, D., Graves, A., Michaelis, J., Li, X., McGuinness, D., et al. (2010). TWC data-gov corpus: incrementally generating linked government data from data.gov. 19th International World Wide Web Conference.
- Ding, L., Lebo, T., Erickson, J. S., DiFranzo, D., Williams, G. T., & Li, X. (2011). TWC LOGD: A portal for linked open government data ecosystems. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 9(3), 1-10.
- Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman, C. (2014). Five stars of linked data vocabulary use. *Semantic Web*, IOS Press, 1-4. Retrieved from <http://www.semantic-web-journal.net/system/files/swj653.pdf/>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Gyeonggi-do Public Data Open Portal (2014, September 15). Retrieved from <http://data.gg.go.kr/>
- Jung, Whan (2013). A study on the method of automatic mash-up considering I/O parameter properties of REST-based open API. Unpublished master's dissertation, Kwangwoon University, Seoul.
- Kang, Du Ho (2011). How public management information system supplied by central government influences on the satisfaction of local government users: Focusing on the sido·saeol PMIS (Busan Metropolitan City). Unpublished master's dissertation, Pukyong National University, Pusan.
- Local Data Open System (LDOS) (2014, October 13). Retrieved from <http://www.gmap.go.kr/tcportal/saeol/SaeolInfo.do/>
- Ministry of Public Administration and Security (2014). Client's guide to build a shared service

of national information resources.

Ministry of Public Administration and Security (2011). Development guidelines of shared services.

Retrieved from

http://www.mogaha.go.kr/cmm/fms/FileDown.do?atchFileId=FILE_00000000023299&fileSn=1/

Ministry of Security and Public Administration (2014). New world made of data. The 1st year seminar of ACT ON PROMOTION OF THE PROVISION AND USE OF PUBLIC DATA. Public Data Strategy Committee.

National Information Society Agency (2014). Data management paradigm strategy. IT & Future Strategy. Report No. 1.

Open Data Portal (2014, September 14). Retrieved from <http://www.data.go.kr/>

Seoul Open Data Square (2014, October 1). Retrieved from <http://data.seoul.go.kr/>