OPEN ACCESS

공공데이터 포털과 개별 포털 간의 데이터 연계방안 연구*

A Study on Data Linkage Between Public Data Portals and Individual Portals

박 진 호 (Jin Ho Park)**
한 상 우 (Sang Woo Han)***

초 록

공공데이터 포털(data.go.kr)은 우리나라 공공데이터 검색과 활용을 위한 통합창구이다. 2021년 행정안전부는 개별 포털 정비방안을 수립한 바 있다. 개별 포털은 공공데이터 외 국내 공공기관들이 구축한 포털을 통칭하는 표현이다. 정비방안에 따라 공공데이터 포털 운영 주체인 한국지능정보사회진흥원은 공공데이터 포털과 개별 포털을 연계하기 위한 운영 및 데이터 연계방안 수립이 필요하다. 이에 본 연구에서는 우리나라 공공데이터 포털, 미국, 영국, 프랑스 데이터 포털 운영현황과 데이터 연계방식을 조사하여 데이터 연계 시 상향식 방법 채택이 효율적임을 제안하였다. 또한, 연계 시 추진해야 하는 구체적인 절차를 데이터 연동 표준방식 결정, 메타데이터 현황분석, 운영 인프라 확충, 데이터 반입 확인, 서비스 개시 5단계로 나누어 제시하였다.

ABSTRACT

The Public Data Portal(data.go.kr) is a gateway for searching and using public data in South Korea. In 2021, the Ministry of Public Administration and Security established individual portal maintenance plans. Individual portals refer to portals built by public institutions in Korea other than the public data portal. According to the maintenance plan, the Korea Intelligence Information Society, the operator of the public data portal, needs to establish operating and data integration plans to link the public data portal and individual portals. In this study, we investigated the current operating status and data integration methods of the public data portal in South Korea, the United States, the United Kingdom, and France, and proposed that the adoption of a top-down approach is efficient when integrating data. In addition, we divided the specific procedures that should be pursued when integrating data into five stages: determination of data integration standard methods, analysis of metadata status, expansion of operating infrastructure, confirmation of data import, and launch of services.

키워드: 공공데이터 포털, 개별 포털, 데이터 포털 연계, 데이터 연계, 데이터 포털 운영방안
Public Data Portal, Individual Portal, Data Portal Connection, Data Connection, Operating Plan for a Data Portal

---

* 이 논문은 2022년 한국지능정보사회진흥원『공공데이터포털-개별 데이터 포털 간의 데이터 연계를 위한 표준화 방안』보고서 일부분을 수정・보완한 것임.
** 한성대학교 도서관정보문화트랙 조교수(jhp@hansung.ac.kr) (제1저자)
*** 광주대학교 문헌정보학과 조교수(swhan@gwangju.ac.kr) (교신저자)
논문접수일자 : 2022년 11월 29일 논문심사일자 : 2022년 11월 29일 게재확정일자 : 2022년 12월 9일
한국비블리아학회지, 33(4): 249-269, 2022. http://dx.doi.org/10.14699/kbiblia.2022.33.4.249
※ Copyright © 2022 Korean Biblia Society for Library and Information Science
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

www.kci.go.kr

1. 서 론

한국지능정보사회진흥원(이하, NIA)이 운영하는 공공데이터 포털[1]은 우리나라 공공기관이 생성 혹은 취득하여 관리하는 공공데이터를 활용할 수 있는 통합창구이다. 이 서비스는 「공공데이터의 제공 및 이용 활성화에 관한 법률(약칭: 공공데이터법), [법률 제17344호]」에 근거하고 있다. 특히, 동 법 제21조 2항에는 공공기관의 장은 특별한 사유 없이 공공데이터의 연계, 제공 등의 협력 요청을 거부할 수 없다고 명시하고 있다.

공공데이터법이 제정된 것은 2013년으로 이후 지속적인 개정 작업이 이루어지면서 현재에 이르고 있다. NIA는 「지능정보화 기본법」 제12조에 따라 공공데이터 제공의 효율성과 활성화 지원을 위해 공공데이터활용지원센터를 설치・운영하고 있다. 데이터와 관련된 법률을 제정하고 전담 기관을 지정하여 서비스를 수행하는 것은 세계적으로 유례를 찾기 어려우며 우리나라가 선도적인 사례를 제공하고 있다고 볼 수 있다.

공공데이터라는 명칭을 사용하고 있지만 사실 공공데이터 포털에서 핵심 데이터 세트는 정부 데이터로 볼 수 있다. 이는 정부가 공공데이터 개방과 활용에 있어서 시발점이라는 의미를 담고 있다. 통상 이런 방식은 열린 정부(Open Government) 운동 중 하나로 볼 수 있다. 열린 정부 운동은 정부 운영 투명성과 시민참여 확대라는 가치를 추구한다. 특히 저작권과 같은 제한조건 없이 모든 사람이 자유롭게 데이터를 활용, 재활용, 재배포할 수 있는 데이터 개방은 핵심 추진사항이다. 데이터 개방은 단순히 투명성 확보를 넘어 데이터를 근간으로 한 새로운 경제 체계에 대한 준비이자 대응이기도 하다. 소위 데이터 경제(data economy)로 대표되는 새로운 경제 질서를 의미한다. 현재 세계는 데이터를 기반으로 새로운 부가가치 창출이 가능한 데이터 중심 정보화 사회로 빠르게 전환하고 있다. 빅 데이터(Big Data), 인공지능(Artificial Intelligence)은 현재 정보화 사회를 대변하는 용어로 볼 수 있다. 이를 실현하기 위해서는 고품질 데이터가 필요하다. 세계 각국 정부는 고품질 데이터 기반 마련을 위해 노력하고 있다. 열린 정부 데이터는 미국, 영국, 캐나다, 한국, 일본 등 주요국을 중심으로 전 세계 개별 국가 단위로 실현되고 있다. 유럽연합 경우에는 집행위원회인 유럽위원회(European Commission) 주도로 유럽연합 내 모든 정부, 정부 관련 단체, 연구소 등의 데이터 검색을 가능하게 하는 유럽연합오픈데이터포털[2]을 운영하고 있다.

우리나라 공공데이터 포털은 열린 정부 데이터, 공공데이터 개방을 위한 핵심 통로로 볼 수 있다. 현재의 추진방식은 공급자(정부) 중심 양적 공급 방식에서 시장(수요자)이 필요로 하는 다양한 형태 데이터를 민관협업으로 제공하는 방향으로 전환하고 있다(김연균, 2021). 이런 추진정책은 2021년 행정안전부 주관 관련 부처 합동 공공데이터전략위원회에서 결의한 5개 안건에 잘 드러나는데 이 안건 중 하나는 개별 데이터 포털 정비 계획이다. 이 안건은 공공

1) https://data.go.kr
2) https://data.europa.eu/en

www.kci.go.kr

데이터 포털과 별개로 공공기관에서 운영하는 포털 중 개별 운영 인정기준을 미충족하거나 효율성 낮은 포털은 통폐합을 추진하고, 불가피하게 별도 운영되는 포털은 개방 데이터가 공공데이터 포털에 연계되도록 조치하는 것이다. 이 조치는 공공데이터 포털을 중심으로 고품질 데이터를 편리하게 활용하기 위한 목적이 있다고 볼 수 있다. 또한, 각 기관이 개별적으로 운영하면서 발생할 수 있는 특정 데이터 고립 문제 해결과 운영상 어려움 해결, 다양한 데이터와 연계 문제해결 등을 내포하고 있다고 볼 수 있다. 행정안전부가 제시하고 있는 문제점 해결을 위한 실현은 NIA가 운영하고 있는 공공데이터 포털을 통해서 이루어진다. 즉, 현재 공공데이터 포털이 갖는 기능이나 운영방식을 유지할지, 수정할지에 대한 방안 수립이 필요하다. 본 논문은 향후 공공데이터 포털과 개별 포털 연계를 위한 방안 수립이 목적이다. 연계를 위해서는 정책 측면, 데이터 측면, 시스템 측면 등 다양한 관점에서 접근과 실행방안 수립이 필요하다. 즉 단기적으로 즉시 실행 가능한 부분과 장기적으로 개선해야 하는 부분, 시스템 개선을 통해 해결해야 하는 부분과 정책적으로 해결해야 하는 부분이 종합적으로 검토되어야 한다는 의미이다. 그러기 위해서는 개별 포털 별 다양한 구축, 운영현황에 대한 분석을 선행해야 한다.

본 연구에서는 이런 중장기적인 접근점을 취하지 않고, 행정안전부 정책에 따라 즉시 통합 혹은 연계가 필요한 상황에서 공공데이터 포털이 운영상 취해야 하는 조치와 데이터 연계 시 선택해야 하는 방향성을 해외 사례를 중심으로 검토하였다. 운영과 데이터 연계 시 선택 방안은 상향식, 하향식 관점으로 제한하여 검토하였다. 상향식은 개별 포털들이 직접 운영, 등록을 수행하는 경우이다. 하향식은 공공데이터 포털이 운영 정책과 데이터 등록을 모두 책임지고 담당하는 경우를 의미한다. 검토 대상 사례는 우리나라를 포함한 미국, 영국, 프랑스 공공데이터 포털이며, 연계 방안 제안 시 추진해야 하는 추진 절차를 함께 제안하였다. 검토 관점을 운영과 데이터 연계로 한정하고 그 방식을 상향식, 하향식으로 한정한 이유는 본 연구가 단기간에 행정안전부 계획을 실현해야 한다는 점이 가장 크다. 또, 연계 후 기관 이용자들은 운영정책이 일반 이용자들은 데이터 세트 검색과 활용이 중요하다는 점을 감안하였기 때문이다.

2. 선행연구

공공데이터 포털과 관련된 선행연구는 공공데이터 포털 자체 현황분석과 개선을 위한 연구와 해당 포털 데이터를 기반으로 한 연구로 구분할 수 있다. 먼저 공공데이터 포털을 대상으로 한 연구를 살펴보면 한희정 외(2020)는 공공데이터 포털 서비스를 시민참여형 포털로 개선하기 위한 방안을 수립하여 제안하였다. 이는 공공데이터 포털이 단순한 데이터 저장소가 아니라 시민참여와 데이터 활용에 가치를 두고 변화하고 있음을 감안한 것이다. 이에 이 연구에서는 공공데이터 포털에서 필요한 시민참여 기능을 규명하고 현재 운영 중인 공공데이터 포털 이용자 서비스 현황을 분석하여 개선방안을 제시하였다. 연구에서 제시한 개선방안 총 5가지로 첫 번째는 위치기반서비스 제공, 두 번

www.kci.go.kr

째는 분석 도구, 시각화 도구, 보조도구 지원, 세 번째는 시민참여 활성화를 위한 공유환경 구성, 네 번째는 크라우드 소싱 기반 위키, RSS, 다양한 커뮤니티 서비스 지원, 다섯 번째는 프로슈머(prosumer) 형태 참여방식 수용이다.

황성욱, 정예용, 김수정, 오효정(2020)은 공공데이터 포털 이용 활성화를 위해 각국에서 운영하고 있는 SNS 현황과 그에 따른 개선방안을 제안하였다. 조사대상은 인도, 미국, 한국으로 계량 분석, 피드백 분석, 시계열 분석, 정보유형분석을 실시하였다. 연구결과에서 제시한 개선방안은 크게 이용자 유입 및 확보와 이용자 만족도 향상 두 가지 측면이다. 먼저 이용자 유입 및 확보에서는 주기적인 이벤트, 공모전 정보 업로드와 업데이트 및 정책뉴스 실시간 게시 및 질의응답 실시, 신규 이용자를 위한 흥미 유발용 콘텐츠 게시를 제시했다. 이용자 만족도 향상 측면에서는 정보 품질관리방안 확보 및 데이터 표준화 선행, 소셜 필터링 및 추천 시스템을 이용한 개인화 정보제공, 행정정보데이터 등 정보 유형 확장 제공을 제시했다.

김동준 외(2021)는 공공데이터 포털에서 개방한 지방자치단체 데이터 활용방안을 제시하였다. 이는 데이터 개방은 확대되고 있지만 활용이 낮다는 점을 개선하기 위한 방법을 제안한 것이다. 특히 공공데이터 포털에서 개방한 지방자치단체 데이터 세트 활용도를 평가하였다. 데이터 활용도는 데이터 조회, 다운로드 수치를 기준으로 하였다. 그 결과 지방자치단체에서 개방한 공공데이터 주제 분류와 규모는 서로 다르지만, 데이터 활용은 전반적으로 미흡한 수준을 보였다. 해결방안으로 풍부한 메타데이터 기술과 공공데이터 포털과 지방자치단체 데이터 포털 연동을 제안하였다.

서형준, 명승환(2015)은 지자체 공공데이터 개방 현황을 분석하고 향후 발전을 위한 구축 전략을 제안하였다. 이는 중앙정부 중심 데이터 개방도 중요하지만 시민과 가까운 접점인 자체의 노력도 중요함을 강조하기 위한 것이다. 이에 이 연구에서는 자자체 공공데이터 포털 담당자의 설문조사와 함께 서울, 부산, 대전, 전라북도, 경기도 강원도 데이터 포털에 대한 분석을 실시하였다. 결과적으로 제안한 지자체의 공공데이터 포털 구축 전략은 지역별 이용사례 제공, 부가가치가 높은 공공데이터 제공, 공공데이터 전담팀 구축, 공공데이터 포털사이트 독자성 확보, 공공데이터 투입효과의 명시 다섯 가지이다.

김지혜, 김수정(2022)은 공공데이터 포털 이용 활성화를 위해 공공데이터 포털에서 제공하는 정보서비스 유형을 정리하고, 이를 기준으로 국내・외 공공데이터 포털 정보서비스 현황을 분석하였다. 분석 결과를 기반으로 향후 국내 공공데이터 포털 정보서비스 개선방안을 제안하였다. 연구에서 제안한 개선사항은 데이터 접근성과 이해도 향상, 분석과 활용을 위한 서비스 구축, 활용촉진을 위한 소통체계 구축이다.

포털 데이터를 기반으로 한 연구를 검토해보면 조재인(2018)은 공공데이터 포털에서 도서관 관련 데이터 분석을 수행한 바 있다. 이 연구에서는 공개한 데이터 수준과 공개 주체에 따른 도서관 관련 데이터 활용 정도 차이를 분석하였다. 분석 결과 국가 및 공공기관이 개방한 데이터로부터는 연구자료 등을 중심으로 4개 주제군집이, 지방자치단체가 공개한 데이터로부터는 소장자료, 평생교육프로그램 등을 중심으로 8개 주제군집을 형성한 것을 확인하였다.

www.kci.go.kr

또한, 국가 및 공공기관이 공개한 데이터가 학교 및 공공도서관 현황을 주로 공개하고 있는 지방자치단체 공개 데이터보다 활용도가 높은 것을 확인하였다. 공공데이터를 활용한 연구는 이 외에도 상당수로 공공데이터 분석을 활용한 공공자전거 입지 선정에 관한 연구(전준헌, 이우진, 2021), 공공데이터 기반 질병 검색 시스템 구축 연구(박순호, 김영길, 2019), 공공데이터 활용 자연재해 피해액 추정 모형 개발 연구(정근채, 2019) 등 다양하다.

선행연구에서 알 수 있는 사항은 공공데이터를 활용하거나 응용하여 유의미한 결과를 도출하기 위한 데이터 기반 연구가 지속적으로 출현하고 있다는 점과 이런 데이터 통합 창구인 공공데이터 포털 개선을 위한 연구가 주를 이룬다는 사실이다. 본 연구와 직접 관련 있는 연구들은 공공데이터 포털 개선에 관한 것으로 주로 공공데이터 포털 서비스 개선에 초점이 맞추어져 있다. 또한, 지자체 공공데이터 포털도 독립적으로 발전해야 하는 대상으로 보고 있다. 본 연구는 이런 지자체를 포함한 다양한 개별 포털과 통합 창구인 공공데이터 포털 간 연계 방안을 고려한다는 점에서 차별성이 있다.

3. 데이터 포털 운영 사례

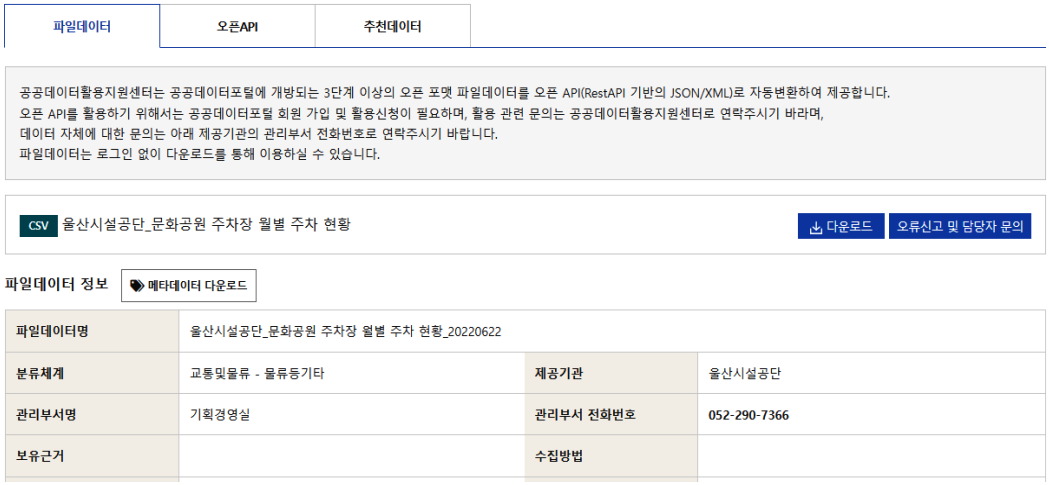
3.1 한국

공공데이터 포털을 통한 데이터 개방은 2013년 「공공데이터의 제공 및 이용 활성화에 관한 법률」 제정・시행과 함께 시작되었다. 이후 공공데이터 포털을 구축하고, 공공기관은 보유하고 있는 공공데이터를 포털에 등록하기 시작했다. 현재(2022년 3월 기준) 공공데이터 포털은 751개 기관, 총 69,057개 데이터를 파일, 오픈 API, 표준데이터 분류별로 등록 및 서비스하고 있다. 공공데이터 포털은 우리나라를 대표하는 공공데이터 주요 공급원으로 다양한 데이터를 무료로 제공하여 공적 영역, 학문 영역은 물론 민간에서 서비스 개발과 사회문제 해결에 기여하고 있다.

현재 공공데이터 포털은 데이터 연계에 있어서 상향식 방식을 취하고 있다고 볼 수 있다. 즉, 등록을 원하는 기관들이 NIA가 운영하는 목록등록관리시스템(all.data.go.kr)에 접속하여 데이터를 직접 등록한다. 등록한 데이터는 담당자 검토와 심의를 거쳐 포털에 발행한다. 각 기관에서 데이터를 등록하기 위해서 각 기관은 공공데이터 목록・데이터 등록기준에 따라 제공 형태별 데이터(파일, 오픈API 등)로 등록 절차를 진행한다. 파일 데이터는 갱신주기(1년 이하)에 맞는 최신 데이터를 제공해야 하며, 오픈 API는 실 데이터 항목과 기술문서가 일치하도록 관리해야 한다. 등록 데이터 서비스 방식은 포털에 직접 등록과 기관링크 제공 모두를 허용하고 있다. 즉 데이터 연계에 있어서 상향식, 하향식 모두를 허용하고 있지만, 공공데이터 관리지침(행정안전부고시 제2021-70호)에 따라 기관링크 제공은 등록자제가 권고사항이다. 사실상 상향식을 지향하고 있다고 보아도 무방하다.

〈그림 1〉은 공공데이터 포털의 검색결과 화면으로 대부분 데이터 세트는 해당 기관이나 서비스로 이동하지 않고 직접 다운로드와 API 연계가 가능하도록 구성하고 있다.

www.kci.go.kr



〈그림 1〉 공공데이터 포털 검색결과 화면

공공데이터 포털의 개방과 이용현황을 보면 2013년 기준년 대비 2022년 현재 공공데이터 개방 건수는 13.098배 증가하였으며, 활용 건수는 2,558배 이상 증가한 수치로 해를 거듭하면서 개방과 활용이 더욱 활발해지는 양상을 보인다(〈표 1〉 참조).

이와 같은 공공데이터 시장 확대는 통합창구인 공공데이터 포털 외에도 기관 자체 데이터 포털 구축 및 데이터 개방 사례가 확대하고 있다는 점에서도 확인할 수 있다. NIA 조사에 따르면 2021년 기준 총 115개 개별 포털 운영을 확인할 수 있다. 115개 포털은 유형별로 데이터 개방 전용, 데이터 분석 플랫폼, 정보공개형 포털로 구분할 수 있다. 대표사례와 함께 유형별 개별 포털을 정리하면 〈표 2〉와 같다.

〈표 2〉에서 확인할 수 있듯이 행정안전부에서 확인한 국내 개별 포털은 총 115개이다. 물론 개별 포털 별로 대상 지식 분야가 모두 다르고, 데이터 세트 형식 다양성 등이 존재한다는 점은 주지할 필요가 있다. 또한, 개별 포털 운영은 현재 법률을 위배하지 않는다. 공공데이터 관리지침(행정안전부 고시 제2021-70호)에는 개별 포털 등 구축·허용기준을 마련하고 있다. 이를 요약하여 정리하면 〈표 3〉과 같다.

예외 규정이 존재하기는 하지만 이용자 입장에서 다양한 공공데이터 검색 및 활용을 위해

〈표 1〉 공공데이터 개방 및 활용 실적

구분		'22년 3월 기준	'13년 기준	비고
개방 건수(누적)		69,057개	5,272개	13.098배↑
데이터 다운로드 및 활용신청 건수(누적)		35,627,297건	13,923건	2,558배↑
오픈API 실시간 정보제공	누적	72,854백만 건	3억 6천만 건	202배↑
	월평균			

(출처: 한국지능정보사회진흥원, 2022)

www.kci.go.kr

〈표 2〉 유형별 개별 포털 현황

구분	개수	대표사례와 특징	책임기관
데이터 개방 전용	39	• 건축데이터 민간개방(open.eais.go.kr) • 건축·주택 인허가 등 전국 지자체 및 산하기관 데이터 통합, 제공(특화성 인정)	국토부
		• 지방행정 인허가 개방(localdata.go.kr) • 지자체 인허가 정보를 시도새올을 통해 통합 제공(특화성 인정)	행안부
		• 나이스 교육 정보 개방(open.neis.go.kr) • 17개 시도교육청 대용량 데이터를 통합 저장소로 수집하여 제공(매일) 등(특화성 인정)	교육부 한국교육학술정보원
데이터 분석 플랫폼	29	• 보건의료 빅 데이터 개방(opendata.hira.or.kr) • 대용량 의료 영상정보, 통계정보 등 제공 • 의료빅데이터 원격분석서비스 제공 등	건강보험심사평가원
		• 환경 빅데이터 플랫폼(www.bigdata-environment.kr) • 과기부 데이터 통합지도 (16개 빅 데이터 플랫폼) 내 환경분야 빅데이터 플랫폼	한국수자원공사
		• 경상남도 빅데이터허브(bigdata.gyeongnam.go.kr) • 경상남도 18개 시군·산하기관 데이터의 수집·연계·관리 및 빅데이터 분석 제공	경남
정보공개형 포털	47	• 대학알리미(www.academyinfo.go.kr) • 공시정보, 대학·특성화·키워드별 학과정보 등 대학 주요 정보를 제공 • 여러 공시데이터를 excel로 다운로드 제공	교육부
		• 국세통계포털(tasis.nts.go.kr) • 국세통계자료를 시각화 서비스, 테마별 통계 등 다양한 방식으로 제공 • 각 통계표를 excel로 다운로드 제공	국세청
		• 공간정보포털(gis.daejeon.go.kr) • 인구, 학교, 시설물 통계정보를 지도에 표출 • 표출 지도정보를 이미지(png)로 저장, 다운로드	대전광역시

(출처: 행정안전부, 2021)

〈표 3〉 개별 포털 등 구축·운영 허용기준

구분	세부기준		주요 내용
필수 요건	비중복성		• 공공데이터 제공을 주 기능으로 수행하는 포털(이하 '개별 포털 등'이라 함)로 공공데이터 포털과 서비스 기능(데이터 제공, 제공신청, 분쟁조정 등)의 중복성이 없을 것
	연계성		• 공공데이터 포털과 시스템 연계 및 데이터 동기화 수행 방안을 제시할 것
선택 요건	법적 근거성		• 개별법에 자체 개별 포털 등 구축을 명시하고 있는 경우
	특화성	제공 방식	• 기관별 자체 개별 포털에서만 데이터 가공·처리·분석이 가능한 정보를 제공하는 경우 • 여러 기관 대용량 데이터를 통합하여 추출·제공하는 경우
		제공 형태	• 3D 공간정보 등 제공 형태가 특화되어 제공 형태와 데이터 분리가 어려운 경우
	보안성		• 제한 접근환경(원격, 물리적 공간 등)을 통해 민감(개인) 정보를 활용해야 하며, 일부 인가된 접근 권한을 가진 이용자에 한해 원격 접근 허용

(출처: 행정안전부 고시 제2021-70호)

www.kci.go.kr

여러 출처를 탐색해야 하는 어려움이 존재하는 것은 문제 상황으로 볼 수 있다. 이런 문제 상황은 법적으로 공공데이터에 대한 통합검색을 지원해야 하는 공공데이터 포털 입장에서도 마찬가지로 볼 수 있다. 우리나라 사례에서 확인할 수 있는 운영과 데이터 연계방식은 다음과 같이 정리할 수 있다.

- 데이터 연계는 상향식과 하향식을 모두 허용하나 하향식이 주를 이룸
- 운영(정책과 표준)은 하향식 채택

우리나라 공공데이터 포털은 본 연구에서 제안하는 운영과 연계방식을 수행할 대상이다. 행정안정부가 제안한 내용은 현재 운영규정이나 데이터 연계방식에서 벗어난다고 보기 어렵다. 단, 데이터 연계와 운영에 있어서 하향식을 모두 채택하고 있는 점은 일관성 있는 서비스 수행과 표준을 기반으로 안정성을 유지하기 위함으로 볼 수 있다.

3.2 미국

미국 공공데이터 포털[3]은 미국 연방 정부 중심이다. 미국 연방 정부 기관이 보유한 공공데이터를 검색, 활용할 수 있는 국가 공공데이터 포털 서비스이다. 이는 공공기관이 구축한 고부가가치 데이터를 국민에게 개방하고, 데이터에 대한 접근성 향상이 목적이다. 데이터 유형별, 주제별, 제공기관별로 검색 서비스를 제공하고, 데이터 활용에 대한 정보와 이용자 참여를 위한 기능을 제공하고 있다. 서비스 측면에서 다양한 연방 정부와 협업을 알 수 있는 부분은 참여하는 조직별로 별도 페이지가 존재한다는 점이다. 즉 데이터 제공기관별로 별도 페이지가 존재하여 기관별 데이터 검색과 기관 데이터 제공 현황 등을 한눈에 파악할 수 있다.

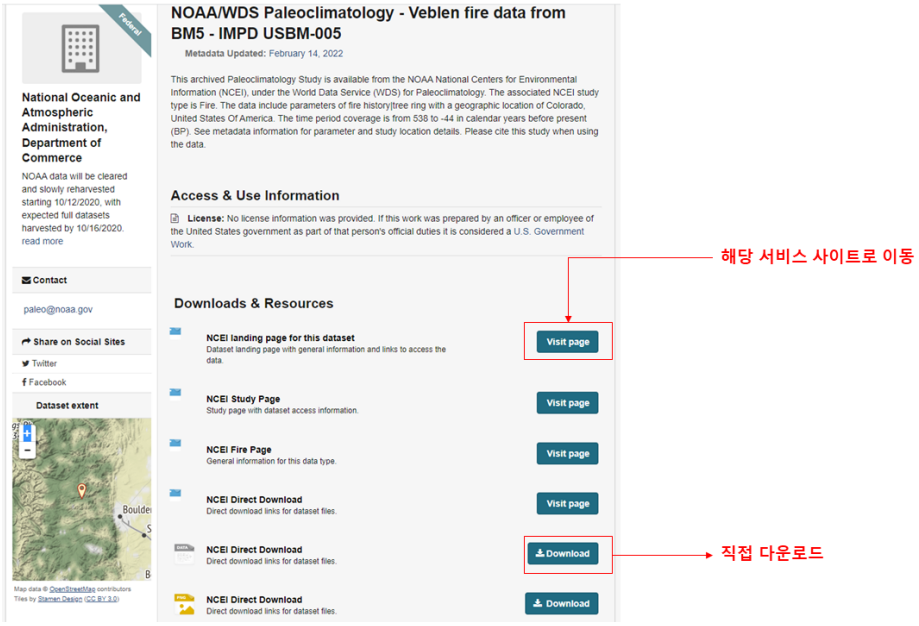
포털 내 개별기관 페이지가 별도로 존재함으로써 참여 기관 간 경쟁을 유도할 수 있으며, 이용자들은 본인이 원하는 기관 데이터에 쉽게 접근할 수 있다. 참여 기관 담당 조직과 담당자 개별 페이지에 대한 정보도 함께 제공해주기 때문에 참여 기관에도 이용자에게도 모두 도움되는 서비스 방식으로 볼 수 있다. 미국 공공데이터 포털 경우는 데이터를 중앙집중형 방식으로 수집하여 제공하고 있으며 각 데이터에 대한 제공기관 접근경로도 함께 제공해주고 있다(〈그림 2〉 참조).

미국이 추진하는 방향은 하향식 방식이 주를 이룬다고 볼 수 있다. 단, 각 데이터에 대한 랜딩 페이지(landing page) 정보도 함께 제공하기 때문에 상향식 접근도 일부 채용하고 있다고 볼 수 있다. 단, 데이터를 제공하는 기관은 해당 데이터를 데이터 포털로 연계하는 작업을 수행해야 하기 때문에 역시 상향식 접근법을 도입했다고 볼 수 있다. 즉, 데이터 연계에 있어서는 상향식, 표준과 정책 수립은 하향식 방식을 도입한 것으로 볼 수 있다.

정책과 표준 관련하여 미국은 두 개 별도 서비스를 운영하고 있다. 하나는 'resources.data.gov'로 이는 데이터 포털에 메뉴로 존재하지만, 별도 서비스 사이트에서 데이터 표준과 데이터 제

3) https://data.gov/

www.kci.go.kr



〈그림 2〉 미국 공공데이터 포털의 검색결과 화면

공자 및 활용자를 위한 지원 정보를 제공한다. 이 서비스는 열린 정부 데이터 법과 연방 데이터 전략 '열린 정부 데이터 법 및 연방 데이터 전략(OPEN Government Data Act and the Federal Data Strategy)'에 기반한다. 이 서비스의 핵심적인 내용은 다음과 같이 정리할 수 있다.

- 데이터 관리 및 거버넌스: 정책구현, 데이터 거버넌스 구조 생성, 연방 정부가 일상적으로 수행할 수 있는 데이터 관리 지원을 위한 안내와 관련 자료 제공
- 데이터 도구: 데이터 실무자 실습 작업을 지원하는 소프트웨어와 관련 자료 제공
- 데이터 인큐베이터: 데이터 관련 프로젝트 양성을 위해 크라우드 소싱, 대회, 민관 연계 지원과 자료 제공
- 기술개발: 데이터 전문가 양성을 위한 안내서와 관련 자료 제공
- 가이드라인: 참고자료로 데이터 관련 문제와 연방 기관에서 적용할 수 있는 다양한 참고자료 제공
- 사례: 연방 정부 데이터 관리자와 데이터 실무자가 수행한 프로젝트 정보 제공

다른 하나는 연방 데이터 전략(Federal Data Strategy)[4]이다. 역시 데이터 포털 메뉴 중 하나이지만 별도 서비스 페이지로 이동하여 서비스를 제공한다. 이는 데이터 기반 사회 변화에 따라 연방 정부가 데이터를 제공, 유지, 사용하

4) https://strategy.data.gov/

www.kci.go.kr

는 방식이 중요하고 데이터 신뢰성을 유지하는 것이 민주적 절차 핵심임을 강조하기 위한 것으로 볼 수 있다. 연방 데이터 전략은 연방 정부가 윤리 거버넌스, 연방 데이터 가치를 최대한 활용할 수 있는 다양한 전략과 지침을 제공하고 있다. 이 두 개 별도 서비스는 데이터 포털이 정책과 표준 측면에서 하향식 방식을 지향하고 있음을 보여준다. 즉, 참여 기관이 준수하고 참고해야 하는 다양한 정보를 통합서비스하고 실제 데이터 연계는 상향식 방식을 유지하고 있다. 단, 상향식 방식을 유지하면서 데이터 보유는 포털과 지방정부가 함께 유지한다는 점이 특징이다. 미국 사례에서 확인할 수 있는 운영과 데이터 연계방식은 다음과 같이 정리할 수 있다.

- 데이터 연계는 상향식 채택
- 운영(정책과 표준)은 하향식 채택

미국은 한국과 마찬가지로 일관성, 안정성 유지를 위해 운영은 하향식을 채택하였다고 볼 수 있다. 반면 실제 데이터를 생산하는 주체가 직접 데이터에 대한 책임을 지고, 포털은 서비스 안정성에 주안점을 두는 연계방식을 유지하고 있는 점이 특징이라 할 수 있다.

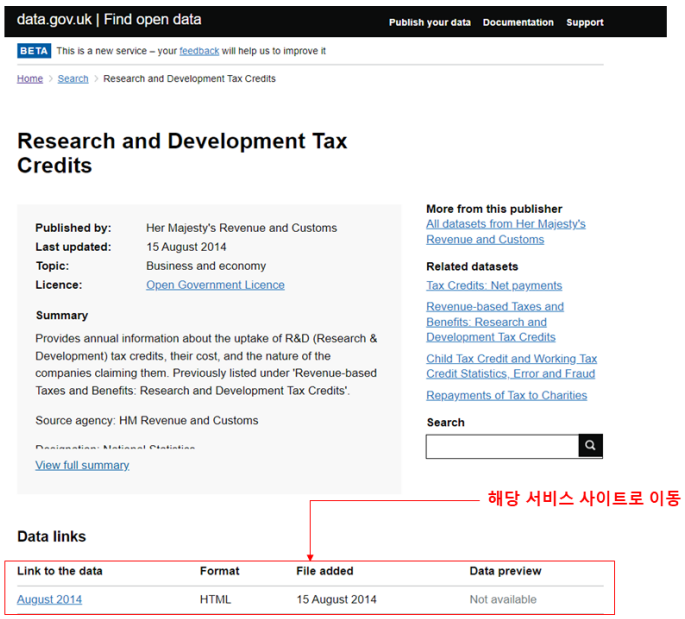
3.3 영국

영국 공공데이터 서비스는 영국 정부 데이터 포털인 'Find open data[5]'가 담당한다. 이 서비스는 영국 정부를 대신해서 국무조정실(Cabinet Office) 산하 GDS(Government Digital Service)가 담당하고 있다. 'Find open data'는 2010년부터 서비스를 시작했다. 지원하는 데이터는 중앙정부, 지방자치단체, 공공기관에서 발행하는 모든 데이터이다. 데이터 검색 후 실제 파일 입수는 해당 서비스 사이트로 연결을 제공하여 해결하고 있다(Find open data, n.d.). 'Find open data'는 영국 공공데이터에 대한 통합검색을 지원하지만, 실제 데이터 관리는 해당 데이터를 생성, 관리, 소유하고 있는 정부 부처와 기관들에서 이루어진다. 〈그림 3〉은 영국 공공데이터 포털 검색결과 화면으로 데이터 획득은 해당 서비스 사이트로 이동하도록 조치하고 있는 것을 확인할 수 있다.

영국 공공데이터 서비스 특징 중 하나는 지방정부 참여에 있다. 'Find open data'에서 지방정부 참여는 이미 이루어지고 있다고 볼 수 있다. 단, 영국은 지방정부연합(Local Government Association, LGA) 참여가 특징이라 할 수 있다. 지방정부연합은 영국 지방정부를 대표하는 범정당 조직으로 지방정부와 관련된 문제에 대해 정치적 아젠다를 형성하고 영향을 주어 국가적 문제에 대한 지역적 해법을 제시하는 것이 목적이다(강혜경 외, 2015). 사실 LGA는 중앙정부에 대한 지방정부 재정확보, 권한 행사 등과 유럽연합 법안에 지방정부 관점을 반영하기 위한 노력 등을 수행하고 지역 중심 과제해결을 위한 협력을 도모하는 것이 주 역할이라 할 수 있다. 최근 LGA는 데이터와 투명성을 중요한 추진과제로 선정한 바 있다. 이는 데이터 개방이 공공서비스를 개선하고 혁신과

5) https://www.data.gov.uk/about

www.kci.go.kr



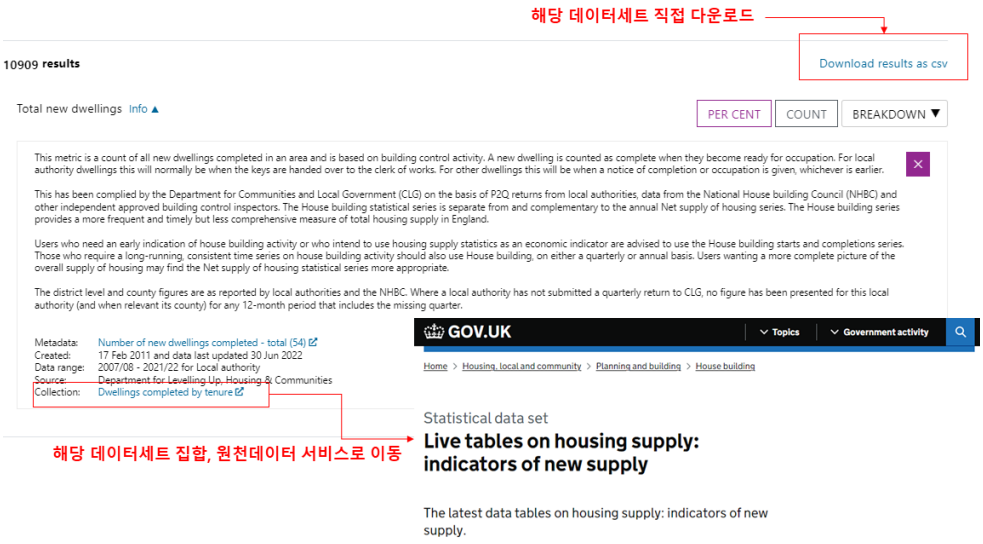
〈그림 3〉 영국 공공데이터 포털 검색결과 화면

성장을 촉진하는데, 도움이 될 수 있음을 인지한 이유이다. Local Government Association (n.d.)는 이를 위해 데이터 개방을 위한 지침, 표준 개발, 모범사례 발굴 등 다양한 노력을 기울이고 있다. LGA는 개방과 투명성 확보가 지방정부 전반에서 공공서비스 재배치를 위한 변화에 반드시 필요한 과정으로 보고 있다. LGA는 데이터 탐색과 시각화를 위해 'LG Inform'[6]이라는 별도 서비스를 운영하고 있다. 해당 서비스에서는 지방정부 관련 다양한 데이터 시각화와 데이터 세트를 제공하고 있다. 해당 데이터 세트는 직접 다운로드 가능하며, 해당 데이터 세트와 관련 데이터 세트 집합 이용을 위해 원천데이터 서비스로 이동하는 기능도 함께 제공하고 있다(〈그림 4〉 참조).

영국도 미국 사례와 비슷한 운영방식과 데이터 연계방식을 보여준다. 영국은 'Find open data'에서 모든 정부 데이터, 공공데이터 통합검색을 지원한다. 하지만 실제 데이터 세트는 해당 서비스로 이동하여 활용할 수 있도록 하고 있다. 검색과 관리에 필요한 체계는 중앙이 담당하지만, 데이터 세트 운영, 관리, 서비스는 해당 기관에 있다. LGA 사례가 독특하지만, 영국 전체 지방정부를 위한 독립 기관 서비스 사례로 볼 필요가 있다. 이 경우도 미국과 유사하게 데이터 관리를 위한 정책, 표준 등은 LGA가 담당하며, 직접 다운로드와 해당 사이트로 이동 모두를 함께 제공하고 있다. 영국 사례에서 확인할 수 있는 운영과 데이터 연계방식은 다음과 같이 정리할 수 있다.

6) https://lginform.local.gov.uk/

www.kci.go.kr



〈그림 4〉 영국 LGA 데이터 세트 검색결과 화면

- 데이터 연계는 상향식 채택
- 운영(정책과 표준)은 하향식 채택
- 지방정부연합을 중심으로 한 지방정부 데이터 포털 존재

영국은 미국과 동일한 방식을 채택하고 있다. 반면 지방정부들을 대상으로 한 별도 포털이 존재한다는 점이 특징이다.

3.4 프랑스

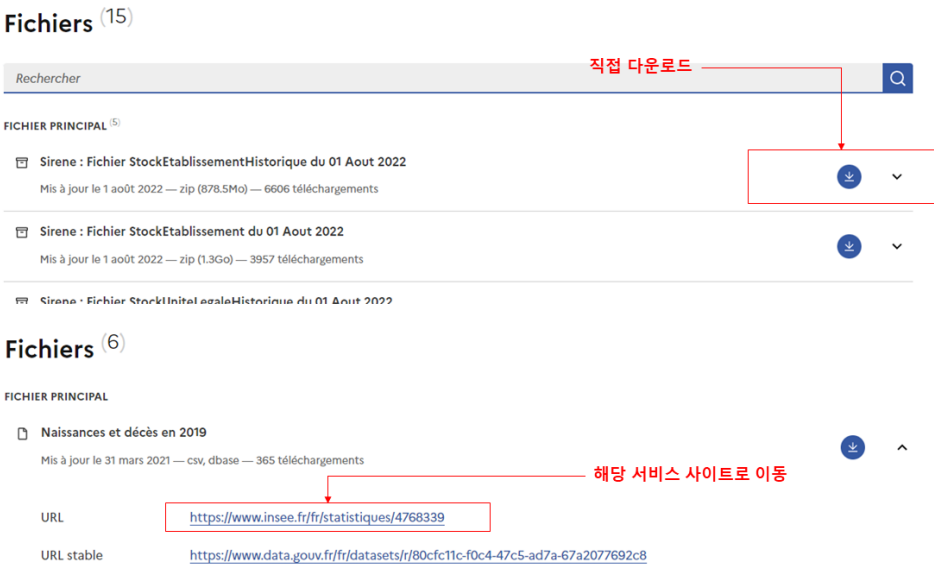
프랑스 공공데이터 포털은 'data.gouv.fr[7]'이다. 프랑스는 유럽 오픈 데이터 서비스 중 가장 질 높은 서비스를 제공하고 있는 것으로 알려져 있다. 이는 2021년 유럽 오픈 데이터 성숙도 보고서(Open Data Maturity Report, 2021)에서 확인할 수 있는데, 프랑스는 유럽연합 27개국 중 가장 높은 순위를 차지하였다.

프랑스는 미국, 영국 사례와 마찬가지로 포털에서는 표준, 정책 수립, 통합검색 중심 서비스를 수행하고 있다. 단, 프랑스는 해당 데이터 세트 검색 시 포털을 통한 직접 다운로드와 해당 서비스로 이동을 혼합하여 사용하고 있다. 〈그림 5〉 윗부분은 공공데이터 포털에서 직접 데이터 다운로드가 가능한 경우를 보여주며, 아래는 해당 서비스로 이동하여 데이터를 활용하도록 하는 모습을 보여 준다.

프랑스 공공데이터 포털의 가장 큰 특징은 최고 데이터 책임조직인 'etalab'을 들 수 있다. 'etalab'은 2019년 10월 30일 법령에 따라 구성한 조직으로 프랑스 데이터 분야 국가 전략 설계와 구현을 조정하는 역할을 담당한다. 구체적으로는 국가 행정부 공공데이터 개방과 공유 정책을 수립하고 공공정보 보급과 재사용 촉진을

7) https://www.data.gouv.fr/

www.kci.go.kr



〈그림 5〉 프랑스 공공데이터 포털 검색결과 화면

위한 정책 수립, 데이터 활용을 위한 데이터 과학, 인공지능 관련 육성과 보급, 정부 투명성과 시민사회 협력을 위한 활동을 수행한다(etalab, n.d.).

Van et al.(2021)에 따르면 프랑스는 장관급으로 최고 데이터 책임자(Chief Data Officer)와 공개 데이터 책임자(Open Data Officer)가 존재한다. Etalab은 최고 데이터 책임자가 이끄는 부서로 영향력이 큰 데이터를 식별하고 부처 간 중개역할도 담당한다. 또한, 프랑스는 8개 기준에 따른 250개 행정 처리 절차와 점수화를 위한 지표를 개발하여 지속적으로 데이터 개방과 서비스 품질을 확인하고 있다. 특히 교통과 같이 중요성이 큰 데이터는 공공데이터 포털 내에 별도 서비스를 개설하여 집중운영하고 있다(교통데이터 'transport.data.gouv.fr'). 이 외에도 프랑스는 이용자 피드백과 모니터링, 소통을 강화하여 서비스를 진행하고 있고, 유럽에서 가장 높은 순 방문자 수를 기록하고 있다. 데이터 세트 통합검색과 재사용을 위해서 가장 중요한 것은 스키마 즉, 데이터 설계서에 있다고 볼 수 있다. 프랑스는 공공데이터 포털에 스키마 검색과 안내를 위한 별도 서비스 체계까지 갖추고 있다. 이 서비스는 'schem.data.gov.fr'[8]이다. 프랑스 사례에서 확인할 수 있는 운영과 데이터 연계방식은 다음과 같이 정리할 수 있다.

- 데이터 연계는 상향식과 하향식 모두 허용
- 운영(정책과 표준)은 하향식 채택

유럽에서 가장 앞선 서비스를 수행하고 있는 프랑스가 취하고 있는 운영방식은 상향식과 하

8) https://schema.data.gouv.fr/

www.kci.go.kr

향식이 조합된 형태로 볼 수 있다. 이는 우리나라와 동일하다. 포털은 운영정책 수립, 표준화 등 하향식 방식을 유지하고 있고, 다른 어떤 국가 서비스보다 체계적인 모습을 보여준다. 데이터 서비스 경우도 직접수집, 저장, 서비스하는 하향식 방식과 해당 기관으로 이동하는 상향식 방식 모두를 취하고 있다.

4. 개선방안

4.1 연계방안

우리나라, 미국, 영국, 프랑스 공공데이터 포털 운영과 데이터 연계방식을 정리하면 〈표 4〉와 같다.

〈표 4〉는 우리나라와 미국, 영국, 프랑스 공공데이터 포털 운영과 연계 방안을 간단히 정리한 것이다. 모든 국가가 공통적으로 운영에 필요한 정책과 표준화 등은 하향식 방식을 유지하고 있다. 즉 정부를 대신해 포털을 운영하는 조직에서 운영정책, 데이터 형식 등을 표준화시켜 운영하고 있다. 반면 데이터 연계와 서비스는 우리나라와 프랑스는 상향식과 하향식을 모두 허용하고 있지만 공통적으로는 상향식을 모두 채택하고 있음을 확인할 수 있다. 즉 운영과 정책은 하향식으로 데이터 연계는 상향식으로 진행하고 있음을 확인할 수 있다.

물론 앞서 검토한 것처럼 우리나라 공공데이터 포털 운영정책은 참여 기관이 데이터를 직접 등록하는 상향식이다. 그러나 현재 공공데이터 포털과 개별 포털을 연계할 때 고려해야 하는 사항은 개별 포털들이 모두 데이터 서비스를 수행하고 있다는 점이다. 어떤 방식이든 장점과 단점을 모두 가지고 있다. 상향식 접근점과 하향식 접근점이 갖는 장・단점은 〈표 5〉와 같이 정리할 수 있다.

〈표 5〉에서처럼 상향식과 하향식 모두 장점과 단점이 뚜렷하다고 할 수 있다. 본 연구에 있어서 중요하게 고려해야 하는 사항 중 하나는 개별 포털들이 공공데이터 포털과 연계 후 서비스를 종료하는 것이 아니고, 유지해야 하고 활성화되어야 한다는 점이다. 또한, 개별 포털과 연계라는 정부 정책 수립과 실현을 위해서는 단기적으로 대규모 예산 투입이 불가피하다. 즉 적정한 예산을 투입하여 현재 안정적인 운영도

〈표 4〉 데이터 포털 운영 연계 방안 정리

구분	주요 내용
우리나라	• 데이터 연계는 상향식과 하향식을 모두 허용하나 하향식이 주를 이룸 • 정책과 표준은 하향식 채택
미국	• 데이터 연계는 상향식 채택 • 정책과 표준은 하향식 채택
영국	• 데이터 연계는 상향식 채택 • 정책과 표준은 하향식 채택 • 지방정부연합을 중심으로 한 지방정부 데이터 포털 운영 사례 존재
프랑스	• 데이터 연계는 상향식과 하향식을 모두 허용 • 정책과 표준은 하향식 채택

www.kci.go.kr

〈표 5〉 데이터 연계 시 상향식, 하향식의 장 · 단점

	상향식	하향식
장점	• 개별기관 보유 데이터 등록이 수월함 • 공공데이터 포털과 개별 데이터 포털 간 연계 체계 구축 유도가 용이함 • 공공데이터 포털에 문제 발생 시 개별 포털을 통해 서비스를 지속할 수 있음 • 다양한 도메인 간 데이터 세트 통합검색이 가능하여 이용자에게 풍부한 정보탐색환경 제공 가능 • 개별 포털이 추구하는 서비스 목표와 서비스 방식, 고유 서비스 유지가 가능함	• 범정부 차원 일관된 메타데이터 항목 및 양질 데이터(활용지원센터 심의 승인) 개방 가능 • 개별기관 운영환경 고려 없이 범정부 차원 일관된 데이터 제공
단점	• 다양한 개별 포털 별 데이터 운영환경과 원천데이터 형태, 반출 가능 형태 모두를 고려해야 함(개별기관 운영환경 등을 고려하여 일부 기관만 연계 가능) - 공공데이터 포털과 목록-데이터 체계, 메타데이터 항목 등이 동일해야 하므로, 연계 이전에 사전 연계 협의 필요 - 개별 포털 및 공공데이터 포털 메타데이터 항목을 매핑시켜야 하기 때문에, 개별 포털에서 추가 정보(메타데이터) 제공 필요 - 개별기관에서 데이터를 유입하는 구조이므로, 공공데이터 포털에서 데이터 심의 반려 시 피드백 제공 어려움 • 개별 포털 고유한 부가 서비스 등을 모두 구현하기 어려움 • 데이터 세트 외에 이용자 피드백 등 추가 연계를 위한 노력 필요	• 공공데이터 포털 선등록에 대한 기관 협조 어려움 예상 • 기관이 보유하고 있지만, 공공데이터 포털에는 미등록된 데이터 발생 가능(정보공개 데이터를 기계판독이 가능 한 형태로 전환 후, 공공데이터 포털 등록을 허용하면 미등록 데이터 발생 가능성을 낮출 수 있음) • 개별 포털들의 데이터 특성을 모두 반영하기 어려움 • 공공데이터 포털에 문제 발생 시 전체 서비스가 중단됨 • 공공데이터 포털이 모든 데이터에 대한 백업과 보존 책임을 가짐으로써 운영인력, 하드웨어 인프라 등 부하가 심함

보장하면서 개별 포털과 원활한 연계까지 고려해야 한다.

이런 고려사항까지 모두 감안하면, 현재 공공데이터 포털과 개별 포털 연계는 상향식 방식을 유지하는 것이 필요하다. 하향식도 불가능한 것은 아니지만 하향식이 갖는 가장 큰 단점은 단기간에 구축이 어렵다는 점이다. 또한, 이미 공공데이터 포털과 개별 포털이 존재하는 상태에서 다시 하나의 보관소와 서비스 시스템으로 데이터를 통합 운영하는 데는 많은 시간을 필요로 한다. 시간뿐만 아니라 기존 서비스가 갖는 이기종 데이터, 운영방안 등을 모두 일원화해야 하기 때문에 안정적인 서비스 운영과 고유 데이터 특성을 모두 반영하기 어렵다. 사실상 데이터 동기화에 있어서 하향식 방안은 구현하기 어렵다. 반면 상향식은 기존 공공데이터 포털과 개별 포털 모두를 유지하면서 전체 데이터에 대한 통합검색을 만들기에 가장 적절한 방법이라 할 수 있다. 단, 이 경우에도 개별 포털 기간 정보 시스템 상황에 따라 다양한 연계방식을 고려해야 할 필요는 존재한다. 즉, 상향식, 하향식 모두 고유 장점과 단점을 갖고 있지만 이미 존재하는 포털 간의 데이터 연계를 고려할 경우에는 상향식 데이터 연계가 가장 적절한 접근점이라 할 수 있다. 단, 상향식으로 데이터를 연동하더라도 장기적으로 표준화와 정책, 교육 등은 하향식 방식을 선택하는 것이 필요하며, 상향식 연동의 경우도 단계별로 추진하는 것이 바람직해 보인다.

www.kci.go.kr

4.2 단계별 추진방안

단기적으로 상향식 데이터 연계를 고려했을 때 필요한 단계별 추진방안은 〈표 6〉과 같이 정리할 수 있다.

〈표 6〉은 단기적으로 추진 가능한 상향식 데이터 연동 방식의 단계별 추진방안이다. 첫 번째 단계는 데이터 연동 표준방식을 결정하는 것이다. 개별 포털 시스템이 존재한다는 것은 이미 기반 시스템이 존재하는 상태이지만 데이터 연계를 위해서는 별도 기술적 조치가 필요하다. 기술적 조치를 위해서 개별 포털은 별도 개발비를 포함한 투자가 필요하다. 데이터 연동에 있어서 고려할 수 있는 방법은 크게 두 가지로 구분할 수 있다. 하나는 API를 활용하고는 것이고 다른 하나는 웹 크롤링이다. 가장 안정적이고 일반적인 방법은 API를 활용하는 것으로 볼 수 있다. 단 이는 기술적인 관점에서 권장 사항으로 개별 포털 상황과 공공데이터 포털 시스템 연계를 위한 예산 현황 등에 따라 다른 방법을 고려해야 할 필요도 있다. 다시 말해 API로 연동을 추진하는 것이 안정적인 방안일 수는 있지만, 개별 포털 상황을 고려하고 빠른 시간에 구축을 목표로 할 경우 달성하기 어려운 상황이 발생할 수 있는 점도 고려할 필요가 있다는 것이다. 때문에, 1차 연동방안은 웹 크롤링도 함께 고려할 필요가 있다. 웹 크롤링과 API 방식을 일반적인 운영환경을 고려해서 간단히 정리 비교하면 〈표 7〉과 같다.

크롤링 경우에도 개별기관 보안 설정 등을 함께 논의하여 공공데이터 포털에서 접근을 허용하는 정책 수립과 실행은 필요하다. 웹 크롤링을 수행할 경우 개발 혹은 별도 상용프로그램을 사용하는 경우에 공통으로 단순히 CSS selector를 사용하기보다는 DOM 기반의 콘텐츠 확장 방안 등을 고려할 필요가 있다. 크롤링 방식이 갖는 가장 큰 단점은 해당 웹사이트 구조가 변경될 경우 해당 코드를 재작성해야 한다는 점이다. 즉 장기 안정성 확보가 어렵다. 또한, 연계하고자 하는 대상 사이트 별로 별도 코드를 개발하고 연계해야 하는 어려움이 존재한다. 구축 시점에서 어려움도 크고, 향후 관리 측면에서도 연계 대상 숫자만큼 어려움이 증가한다고 볼 수 있다. API의 경우도 연계 후 운영 안

〈표 6〉 상향식 데이터 연동의 단계별 추진방안

단계	추진내용
1단계: 데이터 연동 표준방식 결정	• 개별 포털 현황분석을 통해 공공데이터 포털로 연계 가능한 기술적 방식 검토 (예: 웹 크롤링, API 연계 등 가능성 검토) • 각 개별 포털 기술환경 조사
2단계: 메타데이터 현황분석	• 개별 포털별 메타데이터 현황 조사(공공데이터 포털 표준 대비 현황분석) • 연동을 위한 기술적 조치 방안 수립
3단계: 운영 인프라 확충	• 개별 포털 연동을 위한 연동서버, 데이터베이스 서버, 웹 서버 확충 • 개별 포털별 API 설치
4단계: 데이터 반입 확인(시범)	• 시험 서버 등을 통해 개별 포털 데이터 반입 및 결과 확인
5단계: 서비스 개시	• 공공데이터 포털 서비스 개시(공공데이터 포털은 메타데이터와 데이터 세트에 대한 접근 URL 정보만 제공)

www.kci.go.kr

〈표 7〉 웹 크롤링과 API 비교

	웹 크롤링	API
정의	• 웹 사이트 혹은 특정 웹 페이지에서 데이터를 추출하는 프로세스(수집과 분류)	• 응용프로그램에서 사용할 수 있도록 운영체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있는 인터페이스 • 일반적으로 동일 데이터를 사용하는 다른 애플리케이션의 개발을 허용하는 수단
방식	• 수동 수행 혹은 별도 소프트웨어 도구 활용 • 특화된 소프트웨어 활용이 안정적이며, 필요한 부가기능 제공	• API 제공기관에서 별도 개발 및 제공
특징	• 웹의 모든 데이터 추출(개방된 모든 데이터 접근, API보다 많은 종류의 데이터를 확보할 수 가능성이 있음) • 매우 일반적이며 거의 모든 웹 사이트에서 API를 만드는 데 사용할 수 있음 • 데이터를 API로 사용할 수 있는 JSON 파일로 변환할 수 있음 • 해당 웹페이지 html 구조에 맞게 코드를 작성해야 하며, 정제를 위한 노력이 많이 들어갈 수 있음 • 해당 사이트, 페이지 구조가 변경될 경우 코드 수정 불가피 • 해당 사이트 정책에 따라 부하에 따른 차단 가능성 존재	• API는 해당 데이터 세트의 소유자에 의존(운영방식, 요금, 개방 데이터 정보 양 등) • 원하는 데이터에 직접 접근(제공된 데이터에 대한 접근) • 빠른 시간에 정제 형태 데이터 수집 • 데이터 사용에 따른 법적·윤리적 문제가 거의 없음 • 별도 운영조직 등 필요 • 회원가입, 등록 절차가 복잡할 수 있음 • API별로 사용법이 상이하여 별도 노력이 필요할 수 있음

정성 확보나 보안 측면에서 장점이 뚜렷하지만 회원가입, 등록 절차가 복잡할 수 있고, API별로 사용법이 상이하여 별도 노력이 필요할 수 있다는 단점도 존재한다. 사실 크롤링 방식은 운영형태로 보면 하향식에 가깝다고도 볼 수 있다. 그러나 여기서 이를 상향식으로 분류한 이유는 여전히 데이터의 생산 주체는 개별 포털이라는 점 때문이다. 크롤링은 개별 포털 노력을 최소화하고 공공데이터 포털이 단기간에 저비용으로 많은 데이터를 수집하기 위한 최선으로 볼 수 있다. 그러나 본 연구에서 최종적으로 제안하는 데이터 연계 방안은 API 기반 연계이다.

두 번째는 메타데이터 현황분석이다. 각 개별 포털 별로 관리와 서비스에 활용하고 있는 메타데이터 요소를 분석하는 단계로 공공데이터 포털에서 표준으로 사용하고 있는 메타데이터와 차이를 분석한다. 조사 후에는 연동을 위해 필요한 기술적 조치 방안도 함께 수립할 필요가 있다. 크롤링으로 데이터를 수집할 경우 공공데이터 포털 메타데이터를 표준 메타데이터로 활용하면 된다. 만약 API로 데이터 발송한다면 개별 데이터 메타데이터 용어를 그대로 사용하고 공공데이터 포털 관리시스템에서 이를 표준 메타데이터로 변환하는 방식과 개별 포털에서 데이터 발송 시 표준 메타데이터로 변환하는 방식을 함께 고려할 필요가 있다. 장기적으로는 개별 포털 운영 시 공공데이터 포털에서 사용하는 표준 메타데이터로 데이터베이스를 변경하는 방안도 함께 고려할 필요가 있다.

세 번째 단계는 운영 인프라 확충이다. 이는

www.kci.go.kr

연계 대상의 규모와 방식이 결정된 후 기간 인프라인 연동 서버, 데이터베이스 서버, 웹 서버를 확충하는 단계이다.

네 번째는 데이터 반입 확인 단계로 시험 서버를 활용해서 개별 포털 데이터 반입과 결과를 확인하는 단계이다. 결과 확인은 양과 질 모두를 포함해야 한다. 양은 개별 포털과 연계한 데이터가 모두 반입되었는지 여부이며, 질은 표준 형식에 맞는지 여부를 의미한다. 실 서비스 수행 전 안정성을 모두 확인하는 중요한 단계로 볼 수 있다. 여기서 문제가 발생하지 않을 경우, 마지막 단계인 공공데이터 포털에서 서비스를 개시할 수 있다. 최종 연동데이터는 공공데이터 포털에서 메타데이터 정보와 해당 데이터 세트에 대한 접근 URL 정보를 제공하며, 실제 데이터는 개별 포털에서 유지함을 원칙으로 한다. 단, 본 연구에서 제안하는 방안이 향후 현재 형태를 그대로 유지하기 위한 방안은 아니다. 따라서 실제 서비스를 현 단계처럼 추진한다면 이용자에게 향후 서비스 방식, 연계 방식을 변경할 수 있음을 고지할 필요가 있다.

5. 결 론

공공데이터 개방은 단순히 투명한 정부 유지를 위해서만 필요한 것은 아니다. 데이터 세트는 정보화 사회에서 새로운 부가가치를 창출할 수 있는 중요한 자원이다. 대규모 고품질 데이터 세트는 개인보다는 국가가 생산하고 있는 것이 사실이다. 여러 나라 정부들은 공공기관에서 구축한 다양한 데이터를 개방하기 위한 노력을 지속적으로 기울이고 있다. 또한, 대부분 국가들이 해당 공공데이터에 대한 일원화된 접근경로 제공을 위해 데이터 포털을 운영하고 있다. 우리나라도 공공데이터 포털이 존재한다. 데이터에 대한 중요성이 높아지면서 중앙정부뿐만 아니라 하위 정부 부처인 소속 기관, 산하 단체는 물론 지방정부까지 데이터 개방을 위한 노력이 이어지고 있다. 데이터에 대한 공급처가 많아지고 있다는 점은 긍정적이지만 이용자 입장에서 보면 데이터 확보를 위해 다양한 경로를 탐색해야 하는 번거로움이 존재한다. 행정안전부가 개별 데이터 포털 정비방안을 추진한 배경도 이러한 영향이 크다 할 수 있다. 문제는 이를 실현하기 위한 효율적인 방안을 찾는데 있다.

본 연구에서는 공공데이터 포털과 개별 포털이 데이터 연계를 위한 방안을 찾고, 수행절차를 단계별로 제안하였다. 이미 공공데이터 포털은 자체적인 표준 운영방식과 연계방식이 존재하지만 미국, 영국, 프랑스 데이터 포털 연계 방안을 검토하여 한국과 비교하였다. 또한, 개별 포털들이 연계 후에도 안정성을 갖고, 공공데이터 포털에서도 예산문제, 운영 안전성, 단기간 내 구축 등 다양한 현실문제를 해결할 수 있는 방안을 제시하였다. 사례조사결과 우리나라를 포함한 미국, 영국, 프랑스 공공데이터 포털 데이터 연계 방식은 상향식임을 확인하였다. 물론 이는 데이터 연계 시 정책이며 데이터 표준화와 운영에 있어서는 하향식을 채택하고 있음을 확인하였다. 데이터 연계를 상향식으로 추진하고, 실제 구축하기 위해서는 단계적인 접근이 필요하다. 본 연구에서는 이를 5단계로 나누어 데이터 연동 표준방식 결정, 메타데이터 현황분석, 운영 인프라 확충, 데이터 반입 확인, 서비스 개시로 제안하

www.kci.go.kr

였다. 본 연구에서 제안한 방식은 사례조사와 정부 정책 결정에 따른 단기간 사업추진 필요성을 감안한 부분이 가장 크다. 때문에, 장기적으로 다양한 데이터 포털들과 연계에도 적용할 수 있는 유일한 방법이라고 보기는 어렵다. 향후 국내・외 다양한 데이터 정보원으로부터 데이터 반입과 관리, 서비스를 위해서는 다양한 사례연구가 추가로 필요할 것으로 보이며, 다양한 데이터 연계 경험에서 발생하는 문제들과 해결방식을 꾸준히 공유할 필요가 있다.

참 고 문 헌

강혜경, 박세경, 정해식, 김보영, 안혜영, 성은미, 조미형, 주은수, 이민경, 이정은 (2015). 지역사회보장지표 개발 및 지역간 균형발전지원체계 마련 연구. 한국보건사회연구원 정책보고서(2015-104).

공공데이터 관리지침. 행정안전부고시 제2021-70호.

공공데이터의 제공 및 이용 활성화에 관한 법률. 법률 제17344호.

김동준, 김현지, 송채은, 양지우, 김학래 (2021). 공공데이터포털에 공개된 지방자치단체의 공공데이터 활용 방안. 디지털콘텐츠학회논문지, 22(3), 445-452. http://doi.org/10.9728/dcs.2021.22.3.445

김연균 (2021. 10. 18.). 효율 낮은 공공데이터 포털 한 곳에 모은다. 정보통신신문, 출처: https://www.koit.co.kr/news/articleView.html?idxno=89868

김지혜, 김수정 (2022). 국내외 공공데이터 포털 정보서비스 현황 비교 분석. 디지털문화아카이브지, 5(1), 67-83. http://doi.org/10.23089/jdca.2022.5.1.004

박순호, 김영길 (2019). 오픈 소스를 활용한 공공 데이터 기반의 질병 검색 시스템 구현. 한국정보통신학회논문지, 23(11), 1337-1342. http://doi.org/10.6109/jkiice.2019.23.11.1337

서형준, 명승환 (2015). 지자체 공공데이터 개방 현황 및 정책 제언: 공공데이터 포털사이트 운영을 중심으로. 한국지역정보화학회지, 18(4), 1-27. http://doi.org/10.22896/karis.2015.18.4.001

전준헌, 이우진 (2021). 공공 데이터 분석을 통한 서울시 전동 공공자전거의 최적의 입지 선정에 대한 연구. 한국IT정책경영학회 논문지, 13(3), 2475-2481.

정근채 (2019). 공공 데이터를 이용한 회귀분석 기반 자연재해 피해액 추정 모형. Crisisonomy, 15(3), 55-69. http://doi.org/10.14251/crisisonomy.2019.15.3.55.

조재인 (2018). 공공데이터 포털을 통해 개방된 도서관 관련 데이터 분석. 한국비블리아학회지, 29(2), 35-56. http://doi.org/10.14699/kbiblia.2018.29.2.035

지능정보화 기본법. 법률 제18298호.

한국지능정보사회진흥원 (2022). 개별 포털정비 TF 회의 참고자료. 한국지능정보사회진흥원: 대구

www.kci.go.kr

광역시.

한희정, 황성욱, 이정민, 오효정 (2020). 공공데이터포털 이용자 서비스 현황 분석 및 개선방안: 시민참여형 데이터포털을 중심으로. 한국도서관·정보학회지, 51(1), 255-279. http://doi.org/10.16981/kliss.51.1.202003.255

행정안전부 (2021). 공공기관 개별 데이터포털 정비 및 관리계획(안).

황성욱, 정예용, 김수정, 오효정 (2020). 공공데이터포털 이용 활성화를 위한 SNS 운용현황 및 개선방안. 정보관리학회지, 37(2), 23-45. http://doi.org/10.3743/KOSIM.2020.37.2.023

etalab [n.d.]. Mission Etalab. Available: https://www.etalab.gouv.fr/qui-sommes-nous/

Find open data [n.d.]. About Find open data. Available: https://www.data.gov.uk/about

Local Government Association [n.d.]. Data and transparency. Available: https://www.local.gov.uk/our-support/research/data-and-transparency

Van Hesteren, D., Van Knippenberg, L., Weyzen, R., Huyer, E., & Cecconi, G. (2021). Open Data Maturity Report 2021. Luxembourg: Publications Office of the European Union.

• 국문 참고자료의 영어 표기
(English translation / romanization of references originally written in Korean)

Act on Basic Intelligent Informatization. Law No. 18298.

Act on Promotion of Provision and Use of Public Data. Law No. 17344.

Cho, Jane (2018). A study about library-related open data through public data portals. Journal of the Korean Biblia Society for Library and Information Science, 29(2), 35-56. http://doi.org/10.14699/kbiblia.2018.29.2.035

Han, Hui Jeong, Hwang, Sung Wook, Lee, Jung Min, & Oh, Hyo Jung (2020). Analysis of current status and improvement plans of the user service in open data portal: focusing on citizen participation data portal. Journal of Korean Library and Information Science Society, 51(1), 255-279. http://doi.org/10.16981/kliss.51.1.202003.255

Hwang, Sung Wook, Jung, Ye Yong, Kim, Soo Jung, & Oh, Hyo Jung (2020). SNS operation status analysis and improvement plan for facilitating of use of open data portal. Journal of the Korean Society for Information Management, 37(2), 23-45. http://doi.org/10.3743/KOSIM.2020.37.2.023

Jeon, Jun Heon & Lee, Woo Jin (2021). A study on the selection of the best location for electric shared bicycles in Seoul through open data analysis. Journal of The Korea Society of Information Technology Policy & Management, 13(3), 2475-2481.

www.kci.go.kr

Jeong, Keun Chae (2019). Estimation model for damage amounts of natural disasters based on regression analysis using public data. Crisisonomy, 15(3), 55-69. http://doi.org/10.14251/crisisonomy.2019.15.3.55.

Kang, Hye Kyung, Park, Se Kyung, Jeong, Hae Sik, Kim, Bo Young, Ahn, Hye Young, Seong Eun Mi, Jo, Mi Hyeong, Joo, Eun Su, Lee, Min Kyung, & Lee, Jung Eun (2015). A Study on the Development of Local Social Security Indicators and the Establishment of a System to Support Balanced Development Between Regions. Korea Institute for Health and Social Affairs Policy Report (2015-104).

Kim, Dong jun, Kim, Hyeon Ji, Song, Chae eun, Yang, Ji Woo, & Kim, Hak Lae (2021). Methods for utilising local government's public data released to the public data portal. Journal of Digital Contents Society, 22(3), 445-452. http://doi.org/10.9728/dcs.2021.22.3.445

Kim, Ji Hye & Kim, Soo Jung (2022). A comparative analysis on the status of domestic and foreign open data information services. Journal of D-Culture Archives, 5(1), 67-83. http://doi.org/10.23089/jdca.2022.5.1.004

Kim, Yeon Gyun (2021, October 18). Inefficient Public Data Portals Are Collected in One Place. Information and Communication Newspaper, Available: https://www.koit.co.kr/news/articleView.html?idxno=89868

Ministry of Public Administration and Security (2021). Public institution individual data portal maintenance and management plan(draft).

National Information Society Agency (2022). Reference Material for Individual Portal Maintenance TF Meeting. National Information Society Agency: Daegu.

Park, Sun Ho & Kim, Young Gil (2019). Implementation of disease search system based on public data using open source. Journal of the Korea Institute Of Information and Communication Engineering, 23(11), 1337-1342. http://doi.org/10.6109/jkiice.2019.23.11.1337

Public Data Management Guidelines. Ministry of Public Administration and Security Notice No. 2021-70.

Seo, Hyung Jun & Myeong, Seung Hwan (2015). Policy alternatives for public data disclosure by local government: focusing on operation of public data portal site. Journal of Korean Association for Regional Information Society, 18(4), 1-27. http://doi.org/10.22896/karis.2015.18.4.001

www.kci.go.kr

www.kci.go.kr