

개방 공간 데이터의 효과적 활용을 위한 개선 방안에 관한 실증 연구: 입지 분석 적용 사례를 중심으로*

안재성** · 박미라*** · 김화환****

A study on the Improvement of Open Spatial Data for Effective Spatial Analysis Applications: Focused on the Case of Location Evaluation Process*

Jaeseong Ahn** · Mira Park*** · Hwahwan Kim****

요약: 다양한 공간 데이터가 데이터 포털 등의 개방형 서비스를 통해 제공되고 있지만, 표준화 미비와 데이터 중복성 등의 이유로 인해 활용성이 제한되고 있는 실정이다. 다양한 출처의 개방 공간 데이터를 통합적으로 활용할 필요가 있는 사용자는 자료의 선별과 전처리 등 다양한 추가 작업의 부담이 발생하기 때문이다. 본 연구에서는 산업단지 입지분석에 개방 공간 데이터를 활용하는 사례 분석을 통해서 개방 공간데이터 사용 과정의 문제점을 정리하고 효과적인 활용을 위한 개선 방안을 제시하고자 한다. 미래형 산업단지 입지분석을 위한 평가지표를 구축하기 위해서는 개방 공간데이터 검색, 선별 및 전처리 단계에서 다양한 추가 작업이 요구되었고, 동일 주제에 대해서도 상이한 생산 기관들이 중복된 데이터를 제공하여 데이터를 선별하는 작업이 필요하였다. 공간 데이터의 전처리 단계에서는 정확도 및 정밀도 불일치, 객체 지오메트리 불일치, 지오코딩 오류, 좌표계 불일치, 병합 문제 등의 다양한 문제가 발견되었다. 사례 분석에 기초하여 본 연구에서는 개방 공간데이터의 효과적 활용을 위한 개선 방안을 다음과 같이 도출하였다. 첫째, 데이터의 표준화를 통해 효율적 공간 분석을 위한 기반을 강화하여야 한다. 둘째, 중복 데이터 방지를 위해 데이터 생산 주체간의 역할 정립이 필요하다. 셋째, 공간 데이터 개방 서비스와 사용자간 중개 역할을 하는 사용자 중심의 플랫폼 도입이 필요하다.

주요어: 공간 데이터, 개방 데이터, 표준화, 플랫폼

Abstract : A variety of spatial data is provided through open services such as data portals. However, availability is limited due to lack of standardization and data redundancy. Users who need to integrate open spatial data from various sources are burdened with various additional tasks such as sorting and preprocessing of data. In this study, we analyze the problems of using open spatial data through a case study to evaluate location of new industrial complexes, and suggest improvement plans for effective use of open spatial data. In order to construct the evaluation index for the location evaluation of the new industrial complex, various additional works were required in the open spatial data search, sorting and preprocessing stages, and different producers provided the duplicate data for the same subject. In the preprocessing stage of spatial data, various problems such as accuracy and precision mismatch, object geometry mismatch, geocoding error, coordinate system mismatch, and merging problem have been found. Based on the case analysis, this study derives the improvement plan for effective use of open spatial data as follows. First, the basis for efficient spatial analysis should be strengthened through data standardization. Second, to prevent redundant data, it is necessary to establish a role between data producers. Third, it is necessary to introduce a user-oriented platform that acts as a mediator between user and open spatial data services.

Key Words : spatial data, open data, standardization, user-oriented data platform

* 본 논문은 2015년도 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015S1A5A8016346).

** 경일대학교 공간정보공학과 부교수(Associate Professor, Dept. of Geoinformatics Engineering, Kyungil University, jsahn@kiu.ac.kr)

*** 경일대학교 강사(Lecturer, Kyungil University, parkmi0124@naver.com)

**** 교신저자, 전남대학교 지리학과 부교수(Associate Professor, Dept. of Geography, Chonnam National University, h2kim@jnu.ac.kr)

I. 서론

오늘날 대부분의 일상생활은 데이터를 기반으로 한다. 기업, 개인, 센서는 끊임없이 데이터를 만들어 내고 동시에 데이터를 활용한다. IDC(2017)는 일상생활에서 끊임없이 데이터를 활용하는 사회인 데이터 시대(data age)를 예측하기도 하였다. 데이터 사회에서는 사회 전반에 새로운 가치 창출의 수단으로 데이터를 활용한다. 과거에 석유, 전기, 금융자본으로 사회의 경제적 가치를 창출하였듯이, 데이터를 기반으로 새로운 경제적 가치를 만들 수 있는 체계가 구축되는 것이다. Newman(2011)은 데이터 중심의 사회를 토지, 노동, 자본을 뛰어넘어 데이터가 핵심 자원이 되는 데이터 경제(data economy)의 시대로 분석하기도 하였다.

데이터 시대에서 생성되는 다양한 데이터 중 지역, 환경, 보건·위생, 교통, 재난·안전 등의 분야에서 생산하는 공공 데이터가 활용 측면에서 가장 중요한 영역을 차지하고 있다. 우리나라의 「공공데이터의 제공 및 이용 활성화에 관한 법률」에서 공공 데이터는 ‘데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보’를 의미한다. 공공데이터는 재사용에 있어서 최소한의 제약이 있으면서, 무료로 제공되고, 기계가 판독 가능한 형태인 오픈 데이터(open data)의 개념을 강조한다(OECD, 2018). 오늘날 대부분의 국가에서 공공 데이터는 오픈 데이터를 의미하며 개방을 원칙으로 한다.

공공 데이터의 개방은 다양한 긍정적인 효과를 불러온다. 첫째, 공공 데이터 개방은 정부 투명성을 높일 수 있다. 국민들은 정부의 현황이 저장된 개방된 데이터를 통해서 정부의 업무를 쉽게 파악할 수 있다. 둘째, 공공 데이터 개방은 데이터 가공, 분석, 처리, 활용하는 산업을 성장시킨다. 개방된 공공 데이터의 활용이 활성화되면, 데이터를 활용하는 산업뿐 아니라 기존 산업에 데이터를 접목시켜서 새로운 산업이 성장하도록 도움을 줄 수도 있다. 셋째, 공공 데이터의 개방은 시민들의 참여를 활성화시킬 수 있다. 시민들은 개방된 공공 데이터를 활용하여 정부 정책에 의견을 제시할 수 있다. 또한 시민들은 사회 문제를 스스로 해결하는 자발

적 사회 참여 활동에 개방된 공공 데이터를 활용할 수도 있다.

우리나라도 공공 데이터 개방에 따른 긍정적인 효과를 인식하고 다양한 공공 데이터 개방 정책을 수립하고 실행하고 있다. 특히 공공 데이터의 80% 이상이 지리적 요소를 담고 있다는 점을 고려하여 공간 데이터의 개방 및 활용 활성화를 위한 별도의 정책이 실행되고 있다. 2013년부터 2017년까지를 계획 기간으로 하는 제5차 국가 공간정보정책 기본 계획에서 공간 정보의 개방을 정책 목표 중 하나로 제시한 이후 공간정보 개방 정책을 꾸준히 추진하고 있다. 대표적인 사례를 들자면 국가 공간정보 포털을 통한 공간 데이터의 개방 정책을 들 수 있다. 공간 데이터 개방을 통한 산업생태계 활성화를 위해서는 공간 데이터를 유통하고 활용하는 플랫폼이 필요하다. 국가 공간정보 포털은 공간 데이터 유통, 공유, 활용을 위한 플랫폼 역할을 수행하고 있으며, 공간 데이터는 공공 데이터포털과 함께 국가 공간정보 포털에서도 별도로 제공되고 있다.

공간 데이터의 개방이 정부 기관, 기업, 일반 시민에게 긍정적인 효과를 저절로 발생시키는 것은 아니다. 공간 데이터를 개방하여 자유롭게 얻을 수 있다고 하더라도 정부 기관, 기업, 일반 시민들이 원활하게 공간 데이터를 활용할 수 있도록 하기 위해서는 다양한 전문적 기술이 필요한 추가 작업을 수행하여야만 한다(Johnson *et al.*, 2017). Johnson(2016)은 공공 데이터 개방을 위한 비용으로 데이터 수집 비용, 데이터 포맷 변형 작업 비용, 온라인 포털 구축 비용, 데이터셋 유지 관리 비용을 제시한 바 있다. 이러한 비용은 공간 데이터 개방을 위한 준비단계에서 필요한 비용이다. 공간 데이터 개방 이후에 필요한 작업이란 데이터를 제공하거나 활용하는 단계에서 추가로 발생하는 비용이라고 할 수 있다.

개방 공간 데이터를 활용하는 사용자에게는 추가 작업이 데이터 활용에 있어서 중요한 제약 사항이 될 수 있다. 공간 데이터가 개방되더라도 사용자가 활용하지 않는다면 개방에 따른 긍정적인 효과를 전혀 기대할 수 없기 때문에 활용 제약 사항을 줄일 수 있는 방안을 적극 강구해야 할 것이다.

이 연구는 개방된 공간 데이터를 활용하는 단계에서 사용자 입장에서 수행해야 할 추가적인 작업 내용을 분

석하여 개방 공간 데이터 활용 활성화를 위한 제약 조건을 완화할 수 있는 방안을 제시하는 것이 목적이다. 이를 위해서 개방 공간 데이터를 미래형 산업단지 입지 분석에 활용하고, 이 과정에서 발생하는 활용 제약 사항을 실증적으로 분석하였다.

II. 선행연구 검토

개방 공간 데이터에 관한 국내의 선행연구는 크게 두 가지로 분류할 수 있다. 첫 번째는 개방 공간데이터를 활용하는 사례 연구이다. 대표적인 연구를 보면, 개방 공간 데이터를 활용하여 참여지도를 제작하는 방안을 제시하거나(황성수·안재성, 2015), 공간 데이터를 다양한 방식으로 시각화하는 방안을 제시하는 연구(강영욱·김현덕, 2014; 김광섭·이기원, 2015)가 있다. 이러한 연구는 공간 데이터를 활용하여 공간 데이터 제작 및 시각화하는 기법을 개선하는 방안을 제시한다. 특정 분야에 공간 데이터를 활용하는 방안에 대한 사례 연구도 다양하게 이루어지고 있다. 관광 분야(송영선, 2018), 지적 분야(박춘수·홍성언, 2016), 부동산 분야(나성호·김종우, 2019)에 적용 방안을 제시한 연구가 대표적인 사례이다.

두 번째 종류는 개방 데이터의 활용 활성화를 위한 품질 확보 방안에 관한 연구이다. 우리나라의 공공 데이터 품질 확보와 활용 활성화에 관한 연구는 전산 분야와 행정 및 정책학 분야에서 활발히 수행되었다(박고

은·김창재, 2015; 방승열 외, 2015; 김구, 2017; 윤상오·현지우, 2019). 반면, 공간 데이터 품질 확보에 대한 연구는 상대적으로 부족한 실정인데, 이보미·이명근(2017)은 활용도가 높은 국가 지리정보를 선정하고 이 데이터에 대한 품질 관리 절차와 진단 항목을 제시한 바 있다.

이상의 선행 연구 현황을 살펴보면, 개방 공간 데이터 활용 단계에서 나타날 수 있는 문제를 실증적으로 분석한 연구가 없다는 것을 확인할 수 있다. 본 연구는 사용자가 개방 공간 데이터를 활용하는 과정에서 겪게 되는 다양한 문제점을 사례를 통해서 실증적으로 분석하고 제시하였다는 점에서 기존의 연구와는 차별성을 갖는다.

III. 개방 공간데이터 활용 사례 및 활용 데이터

1. 활용 사례

이 연구에서는 개방 공간 데이터를 활용하여 수도권 지역의 미래형 산업단지 후보지 입지를 분석하는 작업을 수행하였다. 미래형 산업단지 후보지 입지 분석에서는 개방 공간 데이터를 활용하여 평가지표에 해당하는 측정지표를 구축하고, 각 측정지표에 전문가 설문조사에서 도출한 가중치를 곱한 값을 합산하여 입지 적합도를 계산한다(그림 1).

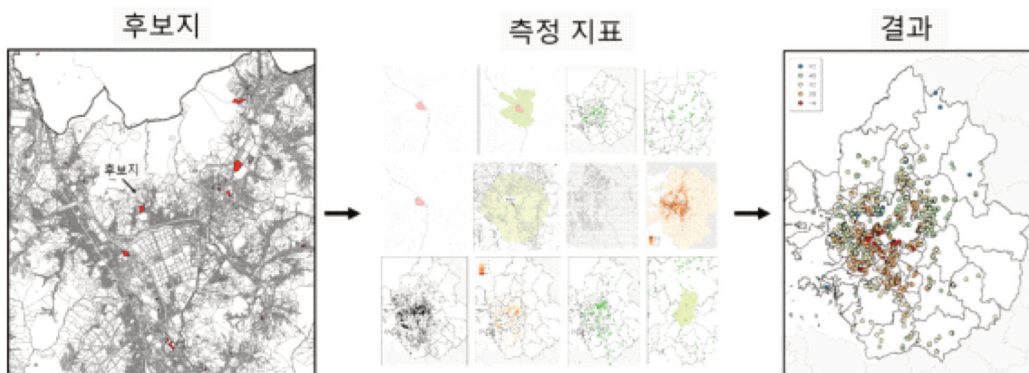


그림 1. 미래형 산업단지 후보지 입지 분석 개요

2. 활용 데이터

지표, 측정지표와 각 측정지표 구축에 사용된 자료 현황은 표 1과 같다.

미래형 산업단지 후보지 입지 분석에 적용한 평가

측정지표 구축에 활용한 데이터는 민간 데이터, 개

표 1. 산업단지 후보지 입지 분석 평가지표, 측정지표 및 활용 자료

우리가 쓸 시간지표랑 비슷

평가지표 체계			측정지표	활용 자료
대	중	소		
산업혁신 여건	지식산업 역량	지식산업 집적도	승용차 10분 거리 내 지식 제조업체·서비스업 체 수	한국SMTP2018(민간자료), FEMIS 자료
			승용차 10분 거리 내 지식 제조업체·서비스업 체 종사자수	한국SMTP2018(민간자료), FEMIS 자료
			승용차 10분 거리 내 기술벤처기업 수	대한민국 벤처산업2017 (민간자료)
		지식산업 성장성	승용차 10분 거리 내 지식 제조업체·서비스업 체의 5년 증가율	한국SMTP2018(민간자료), FEMIS 자료
			승용차 10분 거리 내 지식 제조업체·서비스업 체 종사자수의 5년 증가율	한국SMTP2018(민간자료), FEMIS 자료
			승용차 10분 거리 내 기술벤처기업 수의 5년 증 가율	대한민국 벤처산업2017 (민간자료)
	산업지원 여건	창업지원시설 접근성	승용차 10분 거리 내 기업지원허브, 지식산업 센터, 창업보육센터 수	네이버 지역 API 검색
		기업지원 서비스 접근성	승용차 10분 거리 내 법률·회계·경영·마케팅· 컨설팅·금융·보험 등 기업지원 관련 업체수	한국SMTP2018(민간자료)
			승용차 10분 거리 내 컨벤션센터, 전시판매장 수	네이버 지역 API 검색 결과
	지식자원 잠재력	인적 자원	지자체별 평균 통근시간 내의 생산가능인구 수	국토지리정보원 국토통계자료 (250미터 격자 단위)
			지자체별 평균 통근시간 내의 대학원졸 이상 인구 수(고급노동자)	2010 통계청 센서스 자료 (집계구 단위)
		지식·기술 자원	지자체의 지식재산(특허·실용신안·디자인·상 표) 출원·등록 건수	기초지자체별 출원 등록, 건수(KOSIS)
			승용차 10분 거리 내 대학 입지 여부	개방공공데이터(대학현황)
			승용차 10분 거리 내 연구소(기관) 수	2018 국내 연구소 주소록
도시 어메니티 여건	물류 접근성	-	고속도로 TG와의 네트워크 거리	KTDB(국가교통DB)
		-	주요 간선도로와의 네트워크 거리	KTDB(국가교통DB)
	대중교통 접근성	도시철도 접근성	도보 10분 이내 전철역 유무	서울, 경기, 인천 개방 공공데이터, 도로명주소DB
			전철역과의 보행거리(도보 10분 이내)	서울, 경기, 인천 개방 공공데이터, 도로명주소DB
		버스 접근성	도보 10분 내 버스정류장 수	서울, 경기, 인천 개방 공공데이터, 도로명주소DB
	근무자 어메니티	여가· 위락시설	승용차 10분 거리 내 문화/예술시설(전시관· 미술관·박물관·극장·공연장 등) 수	네이버 지역 API 검색
			승용차 10분 거리 내 쇼핑시설(종합상품판매 업) 수	서울, 경기, 인천 개방 공공데이터 (대규모 점포 인허가 현황)
		건강· 휴식시설	승용차 10분 거리 내 의료시설(병/의원) 수	네이버 지역 API 검색
			승용차 10분 거리 내 스포츠시설(필드·트랙·생 활스포츠시설 등) 수	네이버 지역 API 검색
			도보 10분 거리 내 공원 수	개방 공공데이터(공원 현황)
		기피시설 분포	승용차 10분 거리 내 환경기초시설 수	개방 공공데이터(용도지역지구)

표 1. 계 속

평가지표 체계			측정지표	활용 자료
대	중	소		
부지 적합성 여건	부지 공시지가	-	부지의 제곱미터당 공시지가	개방 공공데이터 (연속지적도, 개별공시지가)
	부지 규모 및 형상	부지 규모	부지의 크기	개방 공공데이터(연속지적도)
		부지 형상	부지의 정형화 여부	-
	주변 지역 갈등유발 여부	-	100m 거리 내 전용주거지역 면적	개방 공공데이터(용도지역지구)
지역발전 효과	지역 경제 가치 증진	-	초등학교 교육환경 보호구역 지정여부	개방 공공데이터(초등학교 현황)
		-	지역의 산업생산·부가가치 유발 효과(지역 생 산·부가가치 대비 증가율)	-
		-	지역의 고용창출 효과(지역 인구 대비 증가율)	-
	지역 사회 가치 증진	-	승용차 10분 거리 내 국가·일반산업단지 유무	개방 공공데이터(산업단지)
		-	지역의 산업구조 개선도	정성
		-	지역의 혁신역량 개선도	정성
		-	지역의 생활·정주환경 개선도	정성

방 공공 데이터로 구분된다. 민간 데이터로는 기업 관련 정보(한국SMTP2018, 대한민국 벤처산업 2017)와 포털 서비스 업체 제공 데이터를 활용하였다. 개방 공공 데이터로는 국가공간정보포털, 각 지자체 공공정보포털, 공공기관 개방 데이터 등을 이용하였다. 연속지적도, 용도지역지구도 등은 국가공간정보포털을 이용해서 구축하였고, 지하철, 버스 정류장 정보 등은 지자체 공공정보포털을 이용해서 구축하였다. 그리고 국토통계자료, 도로명주소 DB 등은 각 기관별 사이트에서 제공하는 데이터를 이용하였다.

바탕순서로 진행하면 됨

IV. 활용 사례 분석 및 개선 방안

1. 분석 단계

이 연구에서는 개방 공간데이터를 활용하여 미래형 산업단지 후보지 입지 분석에 필요한 측정지표를 구축하였다. 측정지표는 공간데이터 형태로 구축되었다. 개

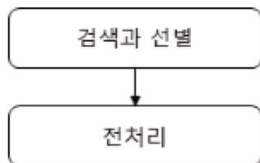


그림 2. 개방 공간 데이터를 활용을 위한 분석 준비 단계

방 공간 데이터를 활용하여 공간 데이터를 구축하기 위해서는 여러 작업 단계를 거쳐야 한다(그림 2). 먼저, 데이터를 검색하고 필요한 데이터를 선별해야 한다. 다음으로 **객체 변환, 지오코딩 등과 같은 전처리 작업**을 수행하여야 한다. 이 연구에서는 각 단계별로 개방 공간 데이터를 활용하는 과정에서 수행한 작업 내용을 분석하였다.

2. 분석 결과

1) 검색과 선별

검색과 선별 단계는 사용자가 필요로 하는 데이터를 검색하고 검색 결과에서 활용할 데이터를 선택하는 작업을 수행한다. 검색 결과에서는 동일 주제이더라도 데이터 생성 기관에 따라서 여러 종류의 데이터를 보여준다. 예를 들어, 국가공간정보포털에서 '시도경계'를 검색어로 검색하면 다양한 기관에서 생성한 데이터가 나열된다. 대표적인 데이터 생성 기관이 통계청과 국토지리정보원이다. 행정경계 데이터는 도로명주소 안내시스템에서도 제공된다.

생성 기관별로 다른 데이터가 검색되면 사용자는 사용 목적에 맞는 데이터를 선택해야 한다. 이 과정이 선별 작업이다. 선별 작업에서 가장 많이 참조하는 데이터가 메타데이터인데, 사용자가 참조한 메타데이터가 활용 측면에서 충분한 정보를 제공해 주지 못하는 경우

최종 갱신시기 정보 누락	
필도	없음
재외파생종	법정종계_시군구
재외파생형	공간
재외파생형	법정종(시군구) 경제도면
갱신주기	변경발생시
최종갱신시기	
구분코드	전체
재외파생종계	Beset:영구(200,000,000,000)
재외파생형	SHA
공간정보분류	국토관리지적정보, 경제

(a) 메타데이터 항목 누락

[illegible]

(b) 사용자 필요 정보 부재

그림 3. 메타데이터 정보 누락 및 필요 정보 부재

가 많이 있다. 그림 3(a)는 최종갱신 시기 정보가 누락되어 있어 데이터의 최신성 여부를 판단할 수 없는 사례이다. 그림 3(b)는 용도지역지구도를 활용하고자 할 때 용도지역 지구 코드에 대한 정보가 메타데이터에 제공되지 않아서 추가로 관련 정보를 찾아야 하는 사례를 보여준다. 메타데이터 입력 정보에 오류가 있거나 사용자가 필요로 하는 정보가 누락되어 있다면 사용자는 추가로 데이터 확인 작업을 수행해야 한다.

2) 전처리

이 연구에서는 정확도 불일치, 정밀도 불일치, 객체 지오메트리 불일치, 지오코딩 오류, 좌표계 불일치, 병합 문제 측면에서의 전처리 작업을 분석한다.

(1) 정확도 불일치

사용자는 정확도 불일치와 관련하여 다양한 추가 작업을 수행하여야 한다. 먼저, 개방 공간 데이터에서 객체 누락 여부를 확인하여야 한다. 개방 공간 데이터에서는 민간 기관의 최신 데이터에 비교했을 때, 공간 객체가 누락되어 있는 경우가 빈번하게 발생한다(그림 4). 이런 경우에 사용자는 다른 데이터를 활용하거나 누락된 객체를 추가하는 작업을 수행하여야 한다.

이 연구의 사례에서는 서로 다른 공간 데이터 간의 연계 작업을 많이 수행한다. 데이터의 갱신주기가 맞지 않거나 속성 항목이 맞지 않으면 데이터를 연계하는 작업을 수행할 수 없는 경우가 자주 발생한다. 예를 들어, 필지 정보가 지가 정보를 연계한다고 가정해 보자. 필

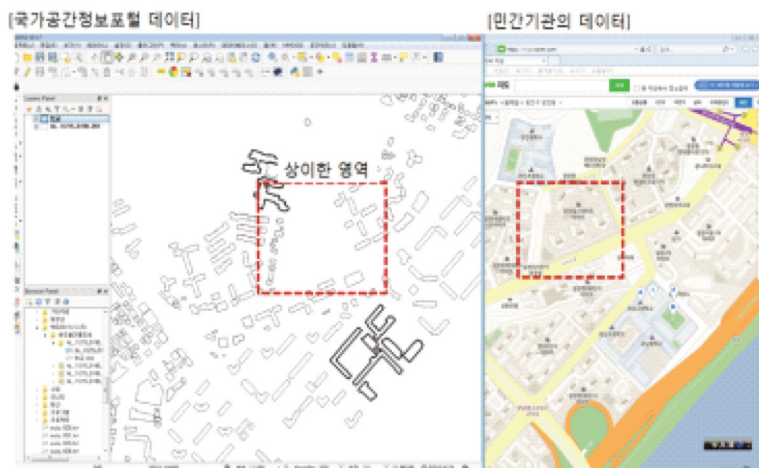


그림 4. 공간 객체의 누락 사례

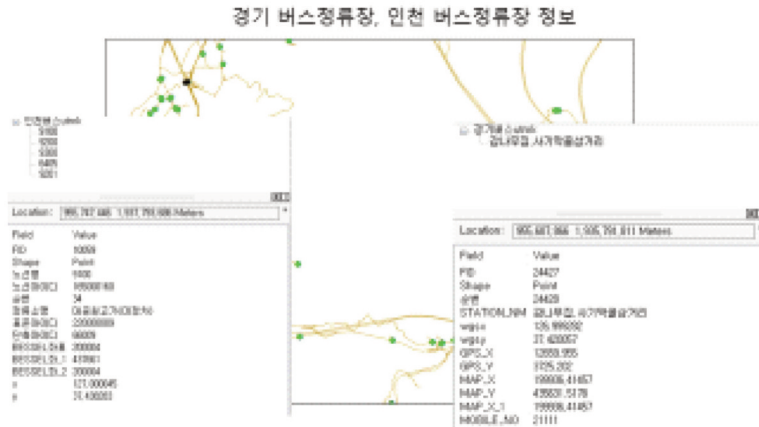


그림 5. 동일 주제이지만 서로 다른 속성 항목을 가지는 사례

지 정보는 2018년 데이터이고 지가 정보는 2019년 데이터인 경우에는 데이터 기준 시점이 맞지 않아서 필지에 연계되지 않는 지가 정보가 다수 발생하게 된다. 건물 정보와 필지 정보를 PNU와 같은 식별자를 이용하여 연계한다고 가정해 보자. 이 경우에도 PNU가 정확하지 않으면 동일한 위치에 있는 건물과 필지 간의 정보가 연계되지 않는 문제가 발생한다.

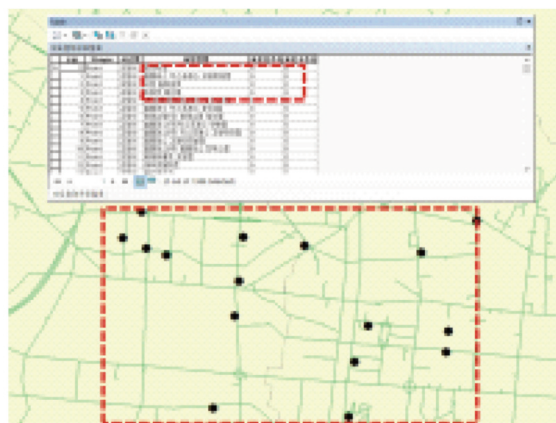
정확도 문제로 인한 데이터 연계의 어려움은 동일한 주제 데이터를 지역별로 다른 속성으로 구축한 경우에도 발생할 수 있다. 버스 정류장 정보의 경우 지자체별로 속성 정보가 다르다. 수도권 지역을 대상으로 하는 경우에는 서울, 경기, 인천의 버스 정류장 정보를 병합해야 한다. 하지만 각 지자체 별로 구축한 버스 정류장

정보의 속성 항목이 다르기 때문에 병합에 문제가 발생할 수 있다(그림 5). 지자체별 버스 정류장 정보를 통합해서 사용하기 위해서는 사용자가 버스 정류장 정보의 속성 항목을 일치시키는 추가 작업을 수행해야 한다.

(2) 정밀도 불일치

일반적인 정밀도의 문제는 공간 데이터의 축척이 다르기 때문에 나타날 수 있다. 이 연구의 사례에서는 행정 자료를 공간 정보로 변환하는 과정에서 정밀도의 문제가 발생하는 것을 확인할 수 있었다.

대부분의 행정 자료는 주소를 기반으로 하며 지오코딩 과정을 거치면서 점형 데이터로 표현된다. 일부 데이터의 경우에는 점형 자료가 분석에 부적합하기도 하다. 점형 데이터로 표현되는 경우에 발생할 수 있는 문제를



- 행정자료의 공간정보 구축 과정에서 정보 제공에 한계가 있음
- 대규모 점포 현황 자료 구축 시 주소 정보를 활용하기 때문에 포인트 정보 구축으로 제한

그림 6. 점형 데이터로 표현되는 대규모 점포 현황 자료

예로 들어 보자. 시장, 백화점, 쇼핑몰과 같은 대규모 점포 현황 자료는 주소로 관리되는 행정 데이터이다. 이러한 자료를 공간 데이터로 변환하기 위해 지오코딩 작업을 거치게 되면 점형 데이터로 표현된다(그림 6). 대규모 점포의 면적을 분석하거나 대규모 점포의 출입구로부터의 거리를 분석하는 경우에는 점형으로 표현되는 데이터를 활용할 수 없다. 이 경우에는 면형 데이터가 적합하다. 사용자가 면형 데이터를 필요로 하는 경우에는 점형으로 변환된 데이터를 다시 면형 데이터로 변환하는 추가 작업을 수행해야 한다.

(3) 객체 지오메트리 불일치

객체 변환 작업은 동일 주제의 데이터일지라도 생성 과정에 따라서 서로 다른 지오메트리를 가질 수 있는 경우에 나타날 수 있는 문제이다. 용도지역지구도의 공원 데이터는 면형 데이터이지만, 지자체에서 관리하는 공원현황 정보는 주소 기반의 행정 정보이기 때문에 지오코딩 작업을 거치게 되면 점형 데이터로 구축된다(그림 7). 이 경우에는 서로 다른 지오메트리를 갖는 공원 자료를 동일한 방식으로 활용할 수 없다. 사용자는 서로 다른 지역의 공원 레이어를 점형 데이터나 면형 데이터로 통일하는 추가 작업을 수행하여야 한다.

(4) 지오코딩 오류

개방 데이터가 주소 정보를 가지게 되면 지오코딩 작업을 거쳐서 공간 데이터를 구축할 수 있다. 지오코딩은 자동화 작업이지만 주소 정보의 부정확성과 불완전성으로 인해서 정확하게 변환된 정보를 생성하지 못하는 경우가 자주 발생한다. 지오코딩 결과에서 나타날 수 있는 대표적인 오류 유형은 동일 지점에 결과 값이 중복되거나 누적되는 결과이다. 주소 정보가 완전하

지 않으면 상위 행정구역 중심점에 주소를 매치시키기 때문에 특정 지점에 다수의 점형 정보가 누적된다. 한 필지에 다수의 주소 정보가 매칭되는 경우에도 동일 지점에 점형 정보가 누적된다. 건물 주소만이 매칭되는 경우에, 동일 건물에 여러 주소가 연결되어 있다면 다수의 점형 정보가 누적된다. 지오코딩 결과에 오류가 있으면 사용자는 주소 정보를 수정하거나 결과 포인트 위치를 정확한 지점으로 이동하는 추가적인 작업을 수행해야 한다.

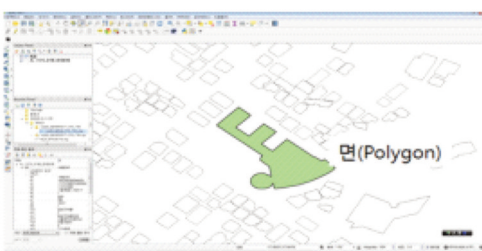
(5) 좌표계 불일치

개방 공간 데이터가 갖는 대표적인 문제점 중에 하나가 좌표계 불일치이다. 좌표계 자체가 다른 경우에는 좌표계 변환 작업을 거쳐서 좌표계를 일치시키면 된다. 문제는 좌표계 정보가 잘못 반영되어 있는 경우에 생긴다. 예를 들어, 데이터의 원 좌표계가 UTM-K이지만 중부원점을 투영원점으로 하는 GRS 80 타원체 기반의 EPSG 5181 좌표계로 메타데이터 정보가 입력되어 있



그림 8. 좌표계 불일치 사례

공원 : 용도지역지구도의 면형데이터



공원 : 공원주소 정보의 지오코딩 변환

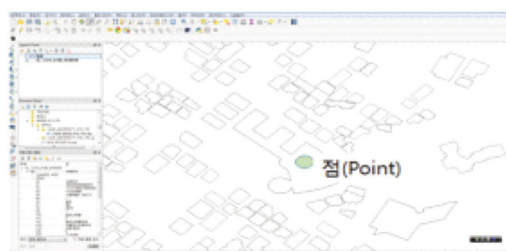


그림 7. 공원 주제 레이어의 객체 지오메트리 불일치



그림 9. 병합으로 인한 경계 불일치

으면, UTM-K 좌표계의 공간 데이터와 함께 활용하는 과정에서 문제가 발생한다(그림 8). 이 문제를 해결하기 위해 사용자는 잘못된 좌표계 정보에 대한 확인 작업을 추가로 수행해야 한다.

(6) 병합 문제

병합 문제는 병합하려는 데이터 생성 기관이 다른 경우에 나타날 수 있는 문제이다. 예를 들어, 시·도·군·구 경계, 읍·면·동 경계 레이어를 중첩해서 사용한다고 가정해 보자. 만약 각 레이어의 생성 기관이 다르다면 중첩하거나 병합해서 사용할 때 경계 불일치 문제가 발생한다(그림 9). 이러한 경우에는 동일 기관에서 생산한 데이터를 사용하거나 각 레이어 경계를 일치시키기 위한 추가 작업을 수행해야 한다.

3. 개선 방안

이 연구에서는 미래형 산업단지 입지분석을 위해 개방 공간 데이터를 활용하는 과정에서 발생할 수 있는 추가적인 작업 내용을 단계별로 분석하고 제시하였다. 각 단계에서 필요로 하는 추가 분석 작업을 최소화할 수 있는 방안은 다음과 같다.

첫째, 개방 공간 데이터의 표준화 작업을 강화하여야 한다. 현재 국내 공간정보 표준 제정 사업이 추진되고 있으나 데이터 구축, 관리, 공유를 포괄하는 전반적인 표준을 모두 제정한 것은 아니다. 공간 데이터 표준이 부족하여 발생하는 문제는 데이터 활용 단계에서도 확인이 가능하였다. 동일 주제의 데이터이지만 속성 항목

이 다르거나, 객체 표현 유형이 다른 경우가 표준화 부재로 인해 나타날 수 있는 대표적인 문제이다. 또한 데이터 생산 기관별로 속성 정확도, 시간 정확도, 정밀도, 내용의 완전성, 갱신 주기 등과 같은 데이터 품질 기준이 다르기 때문에 데이터를 연계해서 활용하는 단계에서 추가적인 작업을 필요로 하는 경우가 많이 발생한다. 사용자의 추가 작업을 최소화하기 위해서는 데이터, 구축, 관리, 공유에 대한 데이터 표준을 제정하여 데이터 생산 기관에서 표준을 준수할 수 있도록 지원해 주는 방안을 마련할 필요가 있다.

둘째, 중복 데이터 방지를 위해 데이터 생산 주체간의 역할 정립이 필요하다. 동일 주제에 대한 중복 데이터의 공개는 사용자의 데이터 검색과 선택을 어렵게 만든다. 또한 표준화되지 않은 중복 데이터의 활용은 좌표계, 경계, 속성 정보 등의 불일치 문제를 발생시켜서 사용자에게 추가적인 작업을 부과시킨다.

셋째, 공간 데이터 개방 서비스와 사용자 간에 중개 역할을 하는 사용자 중심의 플랫폼이 필요하다. 공간 데이터를 개방하는 정부 기관에서 사용자 요구에 부합하는 품질과 내용을 갖춘 데이터를 모두 제공해 줄 수는 없다. 또한 데이터 표준의 부재 및 데이터 생산 주체간의 서로 다른 자료 생산 체계 등으로 인해 데이터 활용을 위한 추가적인 작업은 항상 존재하게 된다. 이러한 문제를 최소화하기 위해서는 사용자의 요구 수준에 부합하는 개방 공간 데이터를 가공하여 제공하는 사용자 중심의 플랫폼이 필요하다. 사용자 중심 플랫폼은

개방 공간데이터를 다양한 사용자 요구에 맞게 가공해서 제공해 주는 역할을 수행해야 한다.

V. 결 론

이 연구에서는 개방 공간 데이터 사용자가 데이터 활용 단계에서 추가로 수행해야 할 작업을 사례를 통해서 분석하였다. 사례 분석에서는 미래형 산업단지 입지분석에 활용하는 측정지표를 구축하기 위해서 다양한 공공 데이터와 민간 데이터를 활용하였으며, 데이터 검색 및 선별과 전처리 단계에서 추가적인 작업을 수행하였다.

데이터 검색과 선별 단계에서는 동일 주제이지만 데이터 생산 기관이 달라서 중복하여 제공되는 데이터를 선별하기 위한 추가 작업이 필요하였다. 전처리 단계에서는 정확도 불일치, 정밀도 불일치, 객체 지오메트리 불일치, 지오코딩 오류, 좌표계 불일치, 병합 문제 등을 해결하기 위한 추가 작업이 필요하였다.

사용자의 추가 작업을 줄여 개방 공간 데이터의 효율적인 활용을 유도하기 위한 개선 방안은 다음과 같다. 첫째, 개방 공간데이터의 표준화 작업을 강화하여야 한다. 둘째, 중복 데이터 방지를 위해 데이터 생산 주체 간의 역할 정립이 필요하다. 셋째, 공간데이터 개방 서비스와 사용자 간에 중개 역할을 하는 사용자 중심의 플랫폼이 필요하다. 향후, 개방 공간데이터 활용을 위한 사용자 중심의 플랫폼을 구현하여 이 연구에서 제시한 개선 방안의 실용성을 검증해 보는 연구가 필요하다.

文 獻

- IDC, 2017, *Data age 2025: The evolution of data to life-critical*, An IDC white paper.
- Johnson, P. A., 2016, Reflecting on the success of open data: How municipal government evaluates their open data programs, *International Journal of E-Planning Research* 5(3): 1-12.
- Johnson, P. A., Sieber, R., Scassa, T., Stephens, M., and Robinson, P., 2017, The Cost of Geospatial Open Data, *Transactions in GIS* 21: 434-445.
- Newman, D., 2011, *How to plan, participate and prosper in the data economy*, Gartner.
- OECD, 2018, *Open government data report: Enhancing policy maturity for sustainable impact*.
- 강영옥·김현덕, 2014, “구글 API를 활용한 공공데이터의 지리적 시각화 연구,” *한국지도학회지* 14(1): 1-15.
- 김구, 2017, “공공데이터 개방정책에 관한 평가적 고찰: 공공데이터포털을 중심으로,” *국가정책연구* 31(2): 57-82.
- 김광섭·이기원, 2015, “전자정부 표준 프레임워크 모바일 실행환경 기반의 공공데이터와 공간데이터 시각화,” *Spatial Information Research* 23(1): 9-17.
- 나성호·김종우, 2019, “공공데이터를 활용한 아파트 매매 가격 결정 모형의 예측 능력 비교: 서울 강남구 지역을 중심으로,” *한국지적정보학회지*, 21(1): 3-12.
- 박고은·김창재, 2015, “공공개방데이터 품질 특성에 관한 연구,” *디지털융복합연구* 13(10): 135-146.
- 박춘수·홍성연, 2016, “지적분야 공공데이터의 이용 활성화 방안,” *한국지적학회지* 32(2): 13-35.
- 방승열·정영기·함병현·김창재, 2015, “공공개방데이터 품질진단모델에 관한 연구,” *한국IT정책경영학회 논문지* 7(4): 23-31.
- 송영선, 2018, “스마트 관광을 위한 개방 공간정보 현황 분석,” *관광레저연구*, 30(11): 05-219.
- 윤상오·현지우, 2019, “공공데이터 개방정책의 실태분석 및 개선방안에 관한 연구: 공공데이터 포털의 국가중점데이터 개방 사례를 중심으로,” *한국공공관리학보*, 33(1): 219-247.
- 이보미·이명근, 2017, “국가지리정보 품질관리 대상 및 방법론 제시,” *국토지리학회지* 51(2): 135-148.
- 황성수·안재성, 2015, “공공정보 개방 활용을 통한 행정 정보와 공간정보의 융합 모색에 관한 연구: 참여 지도 모델링 제안,” *한국정책과학학회보* 19(1): 189-206.

접 수 2019년 04월 20일
최종수정 2019년 05월 27일
게재확정 2019년 05월 31일