

거주민 공간복지 향상을 위한 공공 개방 민원 데이터 분석 모델

- 강동구 공간복지 분석 사례를 중심으로 -

A Public Open Civil Complaint Data Analysis Model to Improve Spatial Welfare for Residents

- A Case Study of Community Welfare Analysis in Gangdong District -

신동윤¹⁾

Shin, Dongyoun¹⁾

Received August 28, 2023; Received September 24, 2023 / Accepted September 25, 2023

ABSTRACT: This study aims to introduce a model for enhancing community well-being through the utilization of public open data. To objectively assess abstract notions of residential satisfaction, text data from complaints is analyzed. By leveraging accessible public data, costs related to data collection are minimized. Initially, relevant text data containing civic complaints is collected and refined by removing extraneous information. This processed data is then combined with meaningful datasets and subjected to topic modeling, a text mining technique. The insights derived are visualized using Geographic Information System (GIS) and Application Programming Interface (API) data. The efficacy of this analytical model was demonstrated in the Godeok/Gangil area. The proposed methodology allows for comprehensive analysis across time, space, and categories. This flexible approach involves incorporating specific public open data as needed, all within the overarching framework.

KEYWORDS: Public Data, Spatial Welfare, Big Data, Text Mining, Complaint Information

키워드: 공공데이터, 공간복지, 빅데이터, 텍스트마이닝, 민원정보

1. 서론

1.1 연구의 배경 및 목적

도시에서 주거환경이란 주거 및 생활 장소를 둘러싸고 있는 생활환경의 총체로서, 좁은 의미로는 주택 주변의 물리적인 환경을 의미하지만, 넓게는 사회적, 경제적, 문화적인 환경을 포함하는 개념이다. 주거환경에 대한 만족은 자신이 사는 사회적, 물리적 환경에 대한 정서적 반응의 결과라고 할 수 있으며, 주거 환경을 구성하는 특정한 변수와의 연관성을 지니고 있다(Shin and Nam, 2012). 특히 산업화와 도시화로 인한 환경의 많은 부분이 바뀌었으며 이러한 요인들은 주민들의 생활환경에 대한 만족도에 많은 영향을 주고 있다 (Kim et al., 2007).

한국 사회는 급속한 산업화와 민주화를 동시에 달성했음에도 삶에 대한 만족도나 행복 수준은 높지 않으며 다양한 사회문제

에 직면해 있으며 이를 과학적으로 파악하기 위해서는 많은 변수를 고려해야 한다(Kim et al., 1998).

이러한 다양한 요소가 영향을 미치는 주거환경 만족도에 대한 평가를 위해 여러 관점에서의 연구가 지속되어 왔다. 일반적으로는 주거환경 만족도란 거주자 개인이 느끼는 거주지의 주관적인 요소와 거주지 주변의 근린 환경에서 얻는 복합적인 만족감으로 정의할 수 있다(Kim and Chung, 2015). 그리고 많은 실증 연구를 통해, 거주민의 설문기반의 주관적 평가로 이루어진 주거환경 만족도는 대표적인 삶의 질 측정지표로 이해되어 왔다 (Kim and Kang, 2019; Kim et al., 2018; Kim and Lee, 2022; Oh and Sim, 2009; Seol and Chael, 2013; Ahn, 2019).

나아가 이러한 주거환경 만족도를 종합하여 도시의 지속가능성과 거주 적합성 여부를 판단하는 지표로 사용하기도 하였다 (Yi and Muhn, 2016).

¹⁾정회원, 단국대 건축학부 조교수 (sindy@dankook.ac.kr)

주거환경 만족도의 결정요인은 많은 연구에서 물리적인 특성과 비물리적인 관점에서 접근하고 있다. 비물리적인 특성은 사회적 요인을 반영한 이웃관계, 갈등, 사회 프로그램 참여 등을 이야기 하고 있다. 그리고 물리적인 특성은 주택 자체의 환경이나 물리적 시설(어린이 놀이터, 주민 회의실, 운동시설, 공원 등 근린생활시설), 그리고 대중교통, 치안 등 접근성 등을 가리킨다(Ahn, 2019). 즉 주거환경 만족도는 주택에 대한 만족도 뿐만 아니라 주변의 도시 환경이 제공하는 복합적인 요소에 대한 총체적 결과로 볼 수 있으며 거주민들의 삶의 질과 동시에 도시의 거주 적합성을 판단하는 중요한 지표로 사용된다(Kim et al., 2016).

1.2 공간복지

이러한 주거환경 만족도의 중요성과 더불어 21세기 ‘복지’의 키워드가 결합되어 최근에는 ‘공간복지’라는 개념으로 만족도를 높이기 위한 요소를 복지의 개념으로 확장되어가고 있다(Kim(1), 2020).

Kim(2), (2020)은 공간복지란 기존의 개인의 임대형 주거 제 공이라는 선별적 복지를 넘어 공공 공간 및 시설을 복지의 요소로 아우르는 장소중심의 복지 개념이라고 설명할 수 있다고 이야기한다. 이는 복지관련 정책이 개인중심에서 지역사회를 기반으로 하는 접근방식으로 변화하는 배경과 맞물려 있다. 일촉에서는 2012년 제정된 사회보장기본법에 따라 사회복지의 개념은 변화를 맞이했다고 설명한다. 저소득 취약계층에 한정적으로 제공되었던 ‘거주복지서비스’ 영역이 일반 국민 모두를 대상으로 하는 사회서비스 개념으로 확장된 것이다. 즉 ‘공간복지’의 화두는 기존의 주거를 위한 경제적 복지 전달체계 중심에서 공공재로서의 공간을 포함하여 그 범위를 확장시킨 것으로 확대되면서 시작되었다고 말하는 것이다(Shin et al., 2014, Je, 2015). 또한 사회보장기본법의 제정에 따라 국민의 소득 및 문화수준 향상에 따라 복지영역 또한 범위가 확대되었다고 한다. 현재 이러한 사회서비스 개념으로서 복지영역은 여러 중앙부서에서 핵심 정책 대상으로 인식하고 있다.

공간복지라는 개념이 많은 곳에서 사용되고 있는 반면 보다 근본적으로 이 개념을 명확하게 정의하고 있는 자료를 찾아보기 힘든 것 또한 사실이다. 따라서, 이 논문에서는 “공간복지란 개인의 삶의 질을 담어내는 주거공간뿐만 아니라 공공을 위한 인프라 공간의 정의”(Yoo et al., 2021)를 기반으로 한다. 나아가 개개인의 상황과 특성마다 공간복지를 다르게 인지함을 인정하고, 이를 배려하는 인프라를 영위할 수 있도록 해야 한다는 것을 포함하도록 한다.

이러한 기초생활 인프라를 지역 단위에서 통합, 연계하여 사회서비스 제공을 도모해야 하는데 서비스에 대한 접근성을 낮추고 동시에 효율성을 증대시키기 위해서 공간복지 개념 도입과

정책의 필요성이 대두된 것이다.

하지만 이러한 계획은 객관적 분석을 통한 결정이 필요하다. 예를 들어, 임의의 계획을 통한 과도한 커뮤니티 시설 확장 등의 운영은 오히려 공간의 낭비와 효율을 저해할 수 있다. 지역별로 커뮤니티 시설을 설치하기 위한 재원은 한정되어 있을뿐더러 관리비에 대한 지속적인 지출 또한 발생하기 때문이다. 더 나아가 시설의 운영이 원활하게 진행되지 못하는 경우 주민들의 분쟁이 유발되어 해당 지역의 가치 하락에 영향을 미치게 된다.

따라서 주거환경 만족도 향상을 위해 객관적 방법론을 정의할 필요가 있다. 본 연구는 그러한 방법이 있어서 근린 환경 수요에 대한 구체적 분석법을 제시하고자 한다. 주거환경 만족도라는 추상적인 개념을 정의하기 위해 공간복지를 구성하는 환경의 인프라를 분석하는 것이다. 이에 활용되는 것이 공공 민원 데이터인데 다른 개방 데이터와 응용하여 실질적인 거주민의 필요에 대한 분석 모델을 제안하고자 한다.

1.3 연구의 방법 및 방향 (공공데이터 활용한 빅데이터 분석환경)

공공 인프라를 포함하는 개념으로써의 공간복지를 실현하기 위해서는 개인들이 요구하는 공공공간의 요구를 파악하고 이해하는 것이 선행되어야 한다. 하지만 복잡한 도시환경에서 이러한 개인들의 요구는 다양한 요소를 고려해야하기에 이를 잘 이해하고 적절한 요소를 계획하는 것은 쉽지 않은 일이다.

따라서, 이러한 계획을 위해 데이터 기반의 다양한 공공데이터의 활용이 증가하고 있다. 특히, 시민들의 생활불편 신고 민원에 대한 분석을 통해 도시문제 해결방안을 강구하고 있다(Kim et al., 2018). 국내에도 서울시의 경우 생활 불편 요인을 신고할 수 있는 시스템인 “서울 스마트 불편 신고 서비스”를 제공하고 이를 바탕으로 근린 환경을 저해하는 도시 문제들을 파악하고 해결하고자 하는 노력을 하고 있다(Kim and Kang, 2019).

이러한 민원데이터는 정형적이지 않으며 확정된 분석방법이 없어 이를 활용하여 새로운 가치를 창출하기에 어려움을 겪고 있으며 이러한 비정형 데이터들을 유의미하게 해석하고 활용할 수 있는 방안에 대한 연구가 최근 빅데이터와 인공지능의 열풍에 타고 진행 중이다. 특히 텍스트 데이터의 경우 개인의 의견을 표출하는 의도로 사용되곤 하는데 텍스트 마이닝 기술을 통한 빅데이터 분석으로 다양한 분야에서 사용되고 있다(Lim and Kim, 2016; Cho and Woo, 2019). 실제로 텍스트 마이닝 기법은 전문가의 주관적 판단이 개입될 가능성을 차단하여 객관적인 결과 도출이 가능하다(Cho and Cho, 2018).

2. 연구 방법

연구 방법을 위해 사전에 고려했던 사항은 다음과 같다.

도시 혹은 거주 만족도와 관련된 연구의 경우 기존의 연구에서 가장 대표적으로 사용한 방법은 설문조사이다. 매우 보편적이며, 주민의 참여를 통하여 현실적인 의견을 수렴할 수 있는 장점이 있으나, 질문 설계의 어려움과 한정적인 표본에 의존하기에 공간적으로 넓은 지역을 분석하는 데에 한계를 가지고 있다(Kim et al., 2020). 또한 분석 대상 선정 및 질문을 위한 사전 조사를 진행해야 한다는 점에서 지속적인 사용이 가능한 방법을 설계하기에는 무리가 있다.

이 외에 Space-Syntax(Hillier et al., 1976) 개념을 활용한 물리적 형태 기반의 분석이 고려되었다. 기존의 현장조사와 같은 데이터 수집의 한계를 극복하기 위한 방법으로는 유효하나, 주거 환경 만족도와 같은 다양한 요소들이 결부된 문제에 대해 유효한 결과를 도출하기 위해서는 물리적 환경 외의 만족도와 직결된 요인을 분석에 포함시키는데 어려움이 있다(Choi and Moon, 2011). 따라서 단순히 인프라에 대한 물리적인 분석만으로는 실효성이 있는 결과를 도출했다고 보기는 어렵다고 판단하였다.

또한 근린시설이나 보행로 등 가시적인 요소들이 잘 설계되어 있다고 하더라도 미처 분석 대상에 포함시키지 못한 요인들이 만족도를 저하시킬 수 있기 때문이다. 따라서 거주자들의 주관에 담긴 데이터를 고려할 수 있는 알고리즘을 제안해야 한다(Jang, 2008).

따라서, 본 연구는 최근에 적극적으로 활용 되고 있는 머신러닝 기술을 이용하여 효과적인 분석과 지속적으로 활용할 수 있는 방법을 고민하였다. 이에, 민원이나 지역과 관련된 뉴스와 같은 주민의 의견을 반영하고는 있으나 데이터의 형식이 비정형 텍스트로 되어 있어 적극적으로 활용되지 못했던 데이터에 집중하였고 머신러닝을 이용하여 객관적인 분석 결과를 도출하는 방법을 제안하고자 한다.

기존의 데이터를 학습기반의 머신러닝을 이용한 분석을 연구 방법으로 제시하고, 민원데이터를 그 분석의 대상으로 하여 공간복지를 위한 프로그램 도출을 연구의 목적으로 한다.

나아가 본 연구는 주거환경 만족도를 객관적으로 분석하기 위한 방법을 제안하고 이러한 분석결과를 근린시설을 포함하는 공간복지 시설에 대한 운영 평가에 활용하기 위한 분석을 포함한다. 이는, 주변 환경에 대한 거주자의 주관적 평가를 객관화하기 위한 방법에 대한 연구를 의미한다.

다시 말해서, 본 연구는 이러한 거주민들의 주거 만족도와 도시문제의 상관관계를 과학적으로 파악하고자, 거주민들의 만족도와 밀접한 연관성을 가지고 있다고 판단되는 민원데이터와의 상관관계 분석을 하였으며 구체적인 분석 방법으로 머신러닝

기반의 데이터 분석을 제시하였다. 나아가 지속적인 연구와 응용분야로의 활용을 위해 범용적인 분석 방법을 고려하고자 한다.

2.1 연구 순서

본 연구의 전체 처리 과정은 Figure 1과 같다. 먼저 데이터 수집은 인터넷 뉴스와 민원 데이터와 같은 비정형 텍스트를 기반으로 한다. 따라서 비정형 데이터를 다루기 위한 전처리 과정을 필요로 하며 객관적인 분석에 장애를 일으킬만한 요소들을 삭제한다. 중복적인 요소들이나 분석 주제와 무관한 내용을 담고 있는 데이터들을 배제하는 것이다. 이렇게 처리된 자료는 토픽 모델링을 활용하여 분석할 수 있다.

그 후에는 커뮤니티 시설 관련 민원만을 추출하여 공간복지 개념과 연계한다. 주거환경 만족도 향상을 위한 방법론을 제시하기 위함이다. 마지막으로 GIS 분석을 통해 데이터를 시각화하여 목적에 맞게 활용한다. 단계별 세부 처리 과정은 3.2절부터 자세히 설명한다.

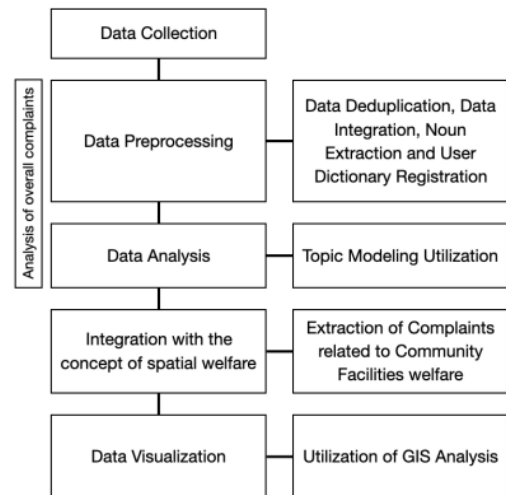


Figure 1. The entire course of study

2.2 데이터 수집

연구의 전반적인 과정을 설명하기 위해 임의로 연구 대상을 서울시로 설정하였고 민원 데이터를 통한 분석과 뉴스 데이터를 통한 분석 두 가지로 진행하였다. 연구에 필요한 자료는 크게 세 가지로 분류할 수 있다. 첫 번째는 거주민들의 만족도에 대한 직접적인 정보를 담고 있는 비정형 텍스트 데이터이다. 민원 데이터의 경우 Table 1에 나타나 있듯이 2015년부터 2020년까지의 통합 민원 데이터 총 82,941건을 수집했다. 추가로 중복되는 내용의 민원을 제거하여 50,985건으로 축소하였다.

두 번째는 유의미한 통계를 위해 비정형 데이터와 결합을 목

적으로 하는 자료이다. Table 1의 민원인 정보 데이터와 조직관리 정보 데이터가 포함된다. 민원인 정보 데이터의 경우 주소 정보를 민원 데이터와 연결하여 민원 내용과 지역의 결합을 위해 수집되었다. 따라서 총 수집된 114,030건의 민원인 정보 데이터 중 주소 정보를 포함하지 않는 데이터를 제외한 48,411건의 데이터를 추출한다. 그리고 조직관리 정보 데이터의 경우 비정형적인 민원 데이터를 어떠한 부서에서 관리하였는지를 통해 데이터 분류 기준을 설정할 수 있기에 수집되었다. 그렇기에 고유의 부서코드 및 부서명을 확보한 총 2,108개의 데이터셋을 확보하였다.

세 번째로는 토지와 건물, 인구와 같은 요소를 담고 있는 공공 개방 데이터이다. 이러한 데이터들은 차후 분석을 더욱 구체화하는 데에 활용할 수 있다. 국토지리정보원에서 제공하는 주민등록인구 31,144명과 유아 인구 29,956명, 고령인구 25,882명과 같은 인구 자료가 수집되었다. 그리고 공시지가 데이터 43,283건과 노후 건축물 14,060개에 대한 정보를 확보하였다. 또한 지역별로 주거환경 만족도를 효율적으로 상승시킬 방안을 제공하기 위해서는 각 데이터와 민원 발생 지역을 접목하는 것이 필수적이기에 행정 경계에 대한 데이터를 확보하였다. 통계청에서 서울특별시 구 분류 데이터 25건과 서울특별시 동 분류 데이터 424건을 수집하였다.

2.3. 데이터 전처리

먼저 수집한 데이터를 확인하고 중복되는 내용의 민원 정보를 제거하는 과정이 필요하다. 민원 정보의 경우 82,941개의 원본데이터에서 전처리를 통하여 최종적으로 50,985건의 데이터로 정리하였다. 민원인 정보 데이터 셋의 경우 114,030건으로 구성된 데이터 중 주소 정보를 보유하지 않은 경우를 제거하여 총 48,411건의 데이터를 확보하였다.

다음으로는 데이터 통합단계가 이루어지는데 각 데이터 셋마다 중복값을 제거함으로써 하나의 분석 데이터 셋으로 통합하였다. 민원 내용 데이터에서 활용할 내부 데이터로 접수 번호(RECEIPT_NO), 민원 제목(APPL_TITL), 부서코드(DPTPR), 민원 내용(APPL_CONT), 민원 접수일시(APPDT)로 설정하고 민원인 정보 데이터에서는 접수 번호(RECEIPT_NO), 고객 번호(CUST_NO), 고객 상세 주소(CUST_ADDR)로 설정한 후 이를 통합하여 접수 번호(RECEIPT_NO), 민원 제목(APPL_TITL), 부서코드(DPTPR), 민원 내용(APPL_CONT), 민원 접수일시(APPDT), 고객 번호(CUST_NO), 고객 상세 주소(CUST_ADDR)로 총 7개의 세부 정보 데이터를 가지고 있는 하나의 데이터 셋을 구성하였다. 최종적으로 데이터 셋 결합 후, 민원 내용과 주소가 확보된 데이터는 총 14,995건이 되었다.

Table 1. Data collection

No	Organization	Data	Data Description	Count
1	Seoul Metropolitan Government Seoul Metropolitan	TNCSB_APPL_REG_INFO	Integrated Complaints Data (2015~2020)	82,941
2		Complainers' Information Data (2012~2020)	114,030	
3		Organization Management Information Data	2,109	
4	National Geographic Information Institute (NGII)	Total Population, Resident Population	Resident Registration Population Data	31,144
5		Infant Population, Elderly Population Data	29,956 / 25,882	
6		Official Land Price Data	43,283	
7	Statistics Korea (Korean Statistical Information Service)	Number of Buildings over 30 Years Old	14,060	
8		Administrative Boundary – Municipal Districts	"Seoul Special City – District Classification"	25
9		"Seoul Special City – Neighborhood Classification"	424	

마지막으로 앞서 진행한 전체 민원 데이터 셋을 활용하여 텍스트 분석을 진행한다. 이를 위해 Komoran 패키지를 활용하여 명사를 추출하였다. 'Komoran'은 텍스트에서 형태소를 분류해 주는 파사어를 말한다. 또한 사용자 사전에서 명사로 등록된 단어가 출현하면 그 단어에 대한 품사를 명사로 정의하는 시스템을 활용할 수 있다. 이러한 사용자 사전 사용은 법적 제도, 건물의 이름, 지명 등과 같이 고유명사를 인식하는 데 활용한다. 동시에 이렇게 추출한 키워드를 확인 후 분석에 유의미한 영향을 주지 못하는 명사에 대한 불용어(Stop word)를 처리하고 고유명사를 확인하여 이 또한 사용자 사전에 등록하도록 한다. 이렇게 명사로 변환된 데이터에서 상위 키워드를 추출하였다(Table 2).

2) 입력 토큰에 내재된 자료 구조를 빌드하고 문법을 검사하는 인터프리터나 컴파일러의 구성 요소 가운데 하나



Figure 3. Visualization of keywords for complaints related to community facilities

3. 대상지 실증 분석

3.1. 고덕/강일을 포함하는 민원 키워드 분석

앞서 제시한 방법을 기반으로 고덕/강일 지구를 대상지를 선정하여 실증 분석 수행하였다. 실증 분석에서는 분석된 데이터를 가시화 하는 방법과 이를 도시 계획에 활용하기 위한 의미를 찾아내기 위한 방법을 포함하고 있다.

먼저 고덕 또는 강일 관련 키워드가 포함된 민원 데이터를 추출하여 현황을 분석하였다. 고덕 및 강일 지구에 대한 내용을 포함하고 있는 민원 내용을 추출하여 1,521건의 데이터를 확보하여, 앞서 진행한 전체 민원 데이터 셋을 활용하여 텍스트 분석을 진행하였다. 그 후 Komoran 패키지를 활용하여 명사를 추출하였다. 또한 사용자 사전 시스템을 활용하여 고유명사를 등록하고, 추출한 키워드를 확인 후 분석에 유의미한 영향을 주지 못하는 명사에 대한 불용어(Stop word)를 처리하였다. 이렇게 명사로 변환된 데이터에서 상위 키워드를 추출함과 동시에 워드클라우드를 활용하여 시각화(Figure 4)할 수 있었다.



Figure 4. Visualization of keywords for complaints related to community facilities in Gangdong-gu

분석으로 확보한 상위 키워드는 아파트, 입주, 단지 등이 나타나며 전체 민원과 유사한 결과가 나타났다(Table 4). 동시에 전체 민원과 달리 ‘명칭’과 같은 키워드가 상위권에 등장했다.

상세 주제는 토픽 분석을 통하여 확인해야 하지만 기존의 토픽과 차이점이 발견되리라 예상된다.

Table 4. Godeok/Gangil area civil complaint top keyword sorting

	Keyword	Count
1	Apartment (아파트)	1669
2	Move-in (입주)	1601
3	Complex (단지)	1591
4	Name (명칭)	1127
5	Construction (공사)	949
6	Civil Complaints (민원)	781

토픽 모델링 결과를 해석 및 시각화하자면 Table 5와 같다. 분석 결과 4개의 토픽이 선택되었다. 토픽 1의 경우 고덕, 강일 아파트 명칭 변경과 관련된 민원이다. 토픽 2의 경우는 고덕 및 강일 아파트 입주 계약 관련 민원이고, 토픽 3은 소득 기준에 따른 아파트 분양 문의이다. 마지막으로 토픽 4는 전반적인 고덕, 강일 지역의 민원 내용을 담고 있다. 전체적으로 전체 민원을 대상으로 한 토픽 분석 결과와 유사하지만 토픽 1과 같이 아파트 명칭 변경에 관한 새로운 토픽이 발견되었다. 또한 4번 토픽의 경우, 전반적인 키워드들이 모두 포함되어 있어 특정 주제로 해석하기에 어려움이 있었다.

Table 5. Godeok/Gangil area civil complaint topic modeling results

Topic 1	Topic 2	Topic 3	Topic 4
Name (명칭)	Kangil (강일)	Godeok (고덕)	Kangil (강일)
Apartment (아파트)	Apartment (아파트)	Income (소득)	Noise (소음)
Occupancy (입주)	Complex (단지)	Gangil (강일)	Complex (단지)
Complex (단지)	Occupancy (입주)	Construction (공사)	Sale of Housing Units (분양)
Prospective Resident (예정자)	Construction (공사)	Occupancy (입주)	Application (신청)
Godeok (고덕)	Godeok (고덕)	Criteria (기준)	Occupancy (입주)
Use (사용)	Contract (계약)	Complex (단지)	Construction (공사)
Change (변경)	Complaint (민원)	Long-term project (장기전)	Housing (주택)
Gangil (강일)	Related (관련)	Sale (분양)	Supply (공급)
Brand (브랜드)	Request (요청)	Installation (설치)	Godeok (고덕)

3.2. 주소정보 활용, 키워드에서 공간으로 확장된 분석

단순한 키워드에서 출발한 분석이, 도시공간과 연계되어 숨은 의미를 찾아내는 과정으로 확장하기 위해서는 공간에 정보를 가시화 하여 상관관계를 직관적으로 가시화 하는 것이 필요하다고 생각하였다. 이를 위해 기존의 민원을 대상으로 14,995개의 데이터 셋 중, 주소 데이터를 포함하는 동시에 커뮤니티 시설명의 키워드를 포함하는 데이터 1,689건을 추출하여 분석에 활용하였다. 그 중에서 흡연, 소음 등 커뮤니티 시설 요구가 아닌 불만사항을 포함한 민원을 제거 후 1,503건의 데이터를 분석의 대상으로 하였다.

본 연구에서 사용될 GIS 데이터로는 행정 경계에 대한 정보를 기반으로 지역의 민원 발생 수, 근린시설이나 교통 현황 등의 데이터를 활용하여 표시하도록 하였다. 따라서 공공 데이터 API와 GIS의 데이터를 결합하여 특정 지역의 민원을 시각적인 이미지로 변환하는 과정을 수행하였다.

커뮤니티 관련 민원들을 공간 정보 파일과 결합하여 구 단위 지역별 매칭을 진행하였다. 전반적인 현황 확인 후, 강동구(강일 및 고덕 지구)를 대상으로 상세 분석을 진행하였다. 서울시 대상 커뮤니티 시설 관련 민원 중 강동구를 대상으로 하는 민원을 추출하여 21건의 데이터를 확보하였고, GIS를 활용한 시각화 및 분석을 위하여 최종 추출된 민원 데이터를 shp파일로 변환하는 과정을 거쳤다. Figure 5는 민원데이터의 구별 발생빈도를 heatmap 방식으로 가시화한 결과이다.

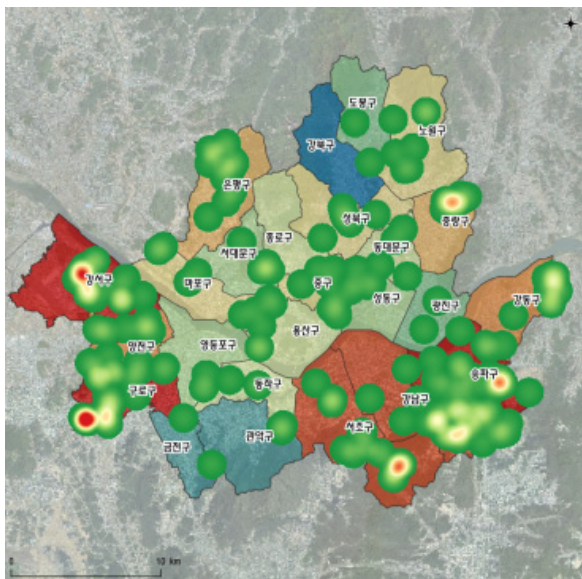


Figure 5. Visualization of civil complaints related to community facilities in Seoul (503 cases)



Figure 6. GIS visualization of civil complaints related to community facilities in Gangdong-gu

민원 내용을 통해 요구 시설을 파악하고 GIS 분석을 통하여 시설 제공의 적절성 및 효과를 분석하였다. 강동구 기준 도출된 21건의 커뮤니티 시설 관련 민원으로 GIS 기반 분석 결과를 토대로 입주민들의 생활환경의 질을 높이는 커뮤니티 시설에 대한 정보를 얻는 것을 가정하고 분석을 진행하였다.

우선 커뮤니티 시설 관련 민원 파악 결과 망월천 육교 설치 요구가 확인되었다. 민원의 위치를 파악하기 위해 가시화한 결과 Figure 6, 7과 같이 발생지역이 강일 리버파크 2단지위에서 다수가 발생한 것을 확인 할 수 있었다.

이를 공간복지 개념에서 해결할 방안을 찾기 위해 주변의 커뮤니티 시설에 대한 정보를 가시화하였다. 망월천 주변의 커뮤니티 시설을 GIS 기반 가시화(Figure 8)를 통해 손쉽게 파악할 수 있었다.

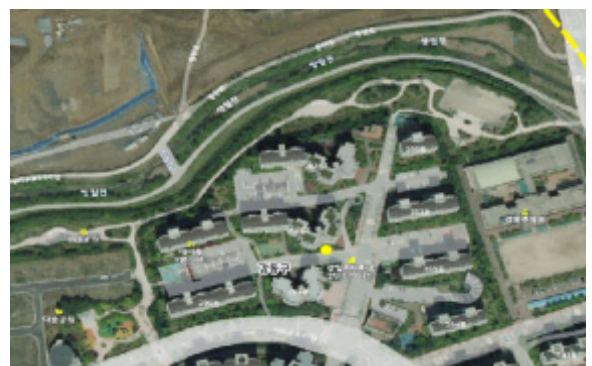


Figure 7. Check the origin of civil complaints through GIS analysis

이를 통하여 현재는 망월천을 사이에 두고 고덕강일 지구와 강일 지구가 분절되어 있는데 이로 인하여 지역 주민들이 주변 커뮤니티 시설에 대한 접근성이 하락하는 공간복지의 저하가

발생함을 알 수 있었다(Figure 9). 즉 망월천에 육교 설치를 통해 이 두 지역, 고덕강일지구 통행이 수월해진다는 상황을 인지할 수 있었다.



Figure 8. Identification of the status of welfare facilities through GIS analysis

이는 단순한 키워드에서 출발한 분석이, 도시공간과 연계되어 숨은 의미를 찾아내는 과정을 보여준다. 공간복지 개념에서 바라본 도시의 문제점을 매우 직관적으로 분석해낸 예시로써, 키워드만으로 분석된 민원의 문제를 도시 공간과 연결하여 분석하는 과정을 보여준다.

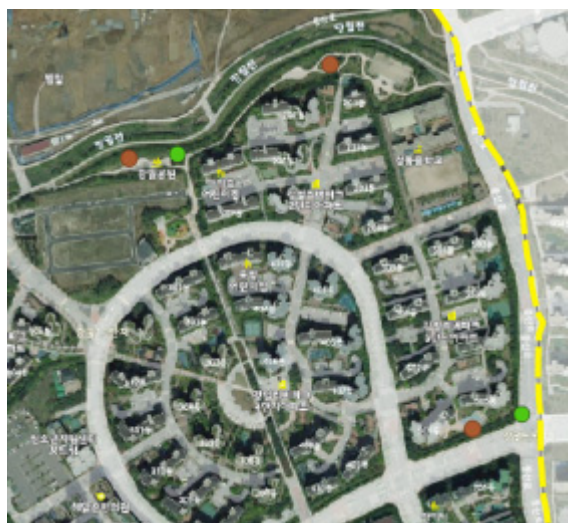


Figure 9. Identify the current state of parks and playgrounds through GIS analysis

4. 결론 및 고찰

본 연구는 위와 같이 빅 데이터 분석을 기반으로 하는 민원 토픽 모델링을 통하여 거주민 니즈 분석 모델을 만들고, 이를 GIS 기반 가시화를 통하여 직관적으로 도시공간으로 확장하여 원인을 파악하고 분석할 수 있는 환경을 구축하는 과정을 보여준다. 강동구를 대상으로 하는 실증 분석을 진행하여 이를 검증하였다.

본 연구는 신뢰성이 높은 공공데이터를 기반으로 다양한 데이터를 분석에 포괄적으로 포함하여 의미 있는 결과를 도출하였다. 나아가 데이터 접근의 일관성을 확보할 수 있는 공공데이터를 활용하여 지속가능한 분석 모델을 만들었다는 것에 의의가 있다. 제시된 방법론에서 전체적인 틀은 유지한 채 필요에 따라 다양한 공공 개방 데이터를 활용한다면 시간, 공간, 대상에 구애되지 않는 포괄적인 분석이 가능하리라 예상된다.

다만 이러한 빅데이터를 이용한 분석은 기존의 통계적 방법의 한계를 극복하기 위한 접근이나 본 연구는 빅데이터를 활용하는 방법에 관한 연구로 제시한 결과 자체를 통계적 접근보다 우수하다고 판단하기에는 데이터의 한계가 있다고 판단한다. 즉 본 연구에서 사용한 데이터의 표본은 정책 대상의 전체 표본이 아닌 여론 및 언론에 표출된 일부를 사용하여 전체의 의견을 수렴한 결과로 보기에는 부족할 수 있다고 여겨진다. 나아가 현재 구축되어 있는 공공 데이터 관리의 한계로 인하여 연구를 진행하는 과정에서 민원인 정보 중 주소 관련 데이터가 부재하는 경우가 잦아 지역별 요구사항 파악에 어려움이 있었다. 주소 데이터가 결측치인 정보를 모두 제외하고 분석을 진행하기 때문에 데이터의 대부분이 사용되지 못하는 상황이 발생하는 것이다. 또한 흡연, 소음, 민원 등 불만 사항에 대한 민원과 시설 요구 민원을 구분함에 어려움도 발견되었다. 배제된 키워드를 포함하고 있는 것과 해당 민원에 시설 요구 내용이 포함되지 않은 것은 상관관계가 불명확하기 때문이다.

이에 대한 해결방안을 마련하고 데이터의 총량이 증가한다면 기존의 목적에 더욱 근접한 연구 결과가 기대된다. 민원 데이터의 내용에 기반한 공간복지 제공 모델을 제시하고 실시간 공공 API 정보를 활용한 GIS 시각화를 통해 보다 즉각적이고 다양한 변수 분석에 범용적으로 사용 가능한 분석법이 될 것으로 기대한다.

References

Ahn, Y. J. (2019). The Effect of Physical Environments in Neighborhood on Residential Satisfaction of Public Central Apartment Housings, *GRI Review*, 21(1), pp. 241–260.
 Cho, J. Y., Cho, K. W. (2018). Topic Modeling on the Adolescent Problem Using Text Mining, *Journal of the Korea*

- Institute of Information and Communication Engineering, 22(12), pp. 1589–1595.
- Cho, K. W., Woo, Y. W. (2019), Topic Modeling on Research Trends of Industry 4.0 Using Text Mining, *Journal of the Korea Institute of Information and Communication Engineering*, 23(7) pp. 764–770.
- Choi, M. O., Moon, Y. S. (2011). An Analysis on the Factors Influencing Busan Metropolitan City Residents' Happiness, *The Korean Journal of Local Government Studies*, 15(1), pp. 277–297.
- Hillier, B., Leaman, A., Stansall, P., Bedford, M. (1976). Space syntax, *Environment and Planning B: Planning and Design*, 3(2), pp. 147–185.
- Jang, H. D. (2008). Factors Affecting Residential Satisfaction and Built Environment Evaluation by Using Them – Focused on Evaluation by Sub-groups Classified Residents' Characteristics on Middle and Small Size Apartment in Seoul, *Journal of the Architectural Institute of Korea*, 24(5), pp. 11–21.
- Je, H. S. (2015). Realization of Space Welfare and Public Architecture for Improving the Quality of Life of the People, *Presidential Commission on Architecture Policy*.
- Kim, B. S., Chung, J. H. (2015). An Analysis on the Determinant Factors of Residential Mobility to Sejong City and Residential Satisfaction, *Korea Real Estate Review*, 25(4), pp. 21–32.
- Kim, C. H., Park H. S., Jung S. M. (2007). A Study on Assessment of Residential Environmental Improvement Projects Achievement, *Journal of Korea Planning Association*, 42(1), pp. 99–112.
- Kim(1), S. Y. (2020). Space is Welfare, *Review of Architecture and Building Science*, 64(4), pp. 15–16.
- Kim(2), H. J. (2020). Space Welfare X Space Doctor, *Review of Architecture and Building Science*, 64(4), pp. 26–29.
- Kim, D. E., Kang, Y. O. (2019). Spatio-temporal Prediction Model of Illegal Parking using LSTM: a Case Study of Civil Complaints in Seoul, *Journal of Korean Society for Geospatial Information Science*, 27(3), pp. 39–47.
- Kim, H. K., Lee, S. W. (2022). The Relationship between Housing Characteristics and Residential Satisfaction in Middle-aged Male One-person Households, *Journal of The Korean Regional Development Association*, 34(2), pp. 165–194.
- Kim, H. S., Im, J. H., Lee, S. S. (1998). A Study on Residential Satisfaction and Preferences of Urban Core Residents, *Journal of the Korean Housing Association*, 9(1), pp. 99–107.
- Kim, J. H., Lee T. H., Ryu, S. E., Kim N. R. (2018). A Study on Text Mining Methods to Analyze Civil Complaints – Structured Association Analysis, *Journal of the Korea Industrial Information Systems Research*, 23(3), pp. 13–24.
- Kim, S. J., Ha, J. H., Lee, S. G. (2020). An Analysis of the Association between Residential Environment Satisfaction and Civil Complaints: Focusing on the Smart Civil Complaints Data in Seoul, Korea, *Journal of Korea Planning Association*, 55(4), pp. 35–49.
- Kim, Y. O. Park, B. N., Kim, G. Y. (2016). Impact Analysis of Residential Environmental Factors on the Residential Housing Satisfaction, *Korea Real Estate Academy Review*, (64), pp. 227–240.
- Lim, M. S., Kim, N. G. (2016). Investigating Dynamic Mutation Process of Issues Using Unstructured Text Analysis, *Journal of Intelligence and Information Systems*, 22(1), pp. 1–18.
- Oh, J. S., Sim Y. M. (2009). A Study on the Management Improvement of Rental Housing for Improving Housing Welfare – Focused on the redevelopment rental housing in Seoul, *Korea Real Estate Society*, 28, pp. 69–87.
- Seol, Y. H., Chael, S. J. (2013). A Study on the Residential Satisfaction of Public Rental Housing in Chungbuk, *Journal of the Korea Academia Industrial Cooperation Society*, 14(9), pp. 4552–4559.
- Shin, E. J., Nam, J. (2012). Determinants of Residential Satisfaction by Residential Environment of Apartment Complexes in Seoul, *Journal of Korea Planning Association*, 47(5), pp. 139–154.
- Shin, Y. G., Yoon, Y. S., Choi, S. M., Yoo, S. H., Kim, H. W., Lee, S. C., Lee, J. K. (2014). An Introduction Plan Research on Station Nursery Facilities applied by 'Spatial Welfare' focused on 'Happy Child Project' by JR East, *Journal of the Korean Society of Industrial Technology*, pp. 264–267.
- Yi, Y. J., Muhn, S. H. (2016). Development and Application of an Apartment Resident Satisfaction Model: Korea Apartment Resident Satisfaction Index (KARSI), *Korean Society of Consumer Studies*, 27.
- Yoo, D. U., Lee, S. J., Park, S. H. (2021). A Study on International Comparison and Improvement Measures for Child-friendly City Indicators for Improving Spatial Welfare of Children, *Journal of the Architectural Institute of Korea*, 37(10), pp. 137–147.