

데이터 플랫폼 활용 영향요인에 대한 탐색적 연구: 한국 공공데이터 포털 사례를 중심으로*

차오신** · 오세환***

요약

정부나 공공기관에서 생성한 데이터를 일반 대중들도 자유롭게 접근, 이용할 수 있도록 공개한 '공공데이터(public data)' 개방 사례는 점차 확대되는 추세에 있으나 미흡한 활용 수준은 문제점으로 지적되고 있다. 데이터 플랫폼 관점에서 공공데이터 활용에 대한 연구가 아직까지 제한적인 가운데 본 연구는 국내 공공데이터 포털에서 제공하는 66,871건의 파일데이터를 대상으로 공공데이터 활용에 영향을 미치는 요인을 탐색적으로 분석하였다. 회귀분석 결과에 따르면 정량적인 데이터셋 설명, 오픈 포맷 형태의 데이터 제공, 사용자 평가, 데이터셋 업데이트 등이 공공데이터 활용에 긍정적인 영향을 미치는 것으로 나타났다. 한편, 데이터셋에 대한 짧은 소개글에서도 다운로드를 발생하지만 설명이 상세할수록 다운로드에 더욱 긍정적인 영향을 미치는 것으로 나타났다.

주제어 : 공공데이터, 공공데이터 포털, 데이터 플랫폼, 데이터 품질, 오픈포맷

An Exploratory Study on the Factors Influencing the Utilization of Data Platforms: Focusing on the Case of the Korean Public Data Portal*

Qiao, Xin** · Oh, Sehwan***

Abstract

Cases of opening "public data"- or data generated by governments or public institutions—so that the general public can freely access and utilize it are gradually increasing, yet the low level of utilization remains a concern. While research on the utilization of public data from the perspective of data platforms is still limited, this study explores the factors influencing public data utilization by analyzing 66,871 file datasets provided by the domestic public data portal. The results of the regression analysis indicate that quantitative dataset descriptions, provision of data in open format, user ratings, and dataset updates have a positive impact on public data utilization. Meanwhile, downloads occur even with brief introductions to datasets, but more detailed descriptions have been found to have a more positive impact on downloads.

Keywords : public data, public data portal, data platform, data quality, open format

Received Oct 21, 2024; Revised Nov 21, 2024; Accepted Nov 25, 2024

* This research was supported by Kyungpook National University Research Fund, 2022. Preliminary findings from this study were presented at the Korea Association of Information Systems (KAIS) 2024 Fall Conference.

** Ph.D. Candidate, School of Business Administration, Kyungpook National University (kyoshin@knu.ac.kr <https://orcid.org/0009-0003-5541-2598>)

*** Corresponding Author, Associate Professor, School of Business Administration, Kyungpook National University (sehwan@knu.ac.kr <https://orcid.org/0000-0001-5823-4588>)

I. 서론

최근 정부나 공공기관에서 생성한 데이터를 일반 대중들도 자유롭게 접근, 이용할 수 있도록 공개한 ‘공공데이터(Public Data)’ 개방 사례가 증가하고 있다. 국내에서는 2013년 「공공데이터의 제공 및 이용 활성화에 관한 법률」(약칭: 공공데이터법)이 시행되면서 공공데이터 포털(Data.go.kr)이라는 데이터 플랫폼을 중심으로 다양한 공공데이터를 Open API 또는 파일 형태로 제공하고 있는 가운데 전세계적으로 40여개 이상의 국가에서 공공데이터 플랫폼이 운영되고 있다(Kubler, et al., 2018; Nikiforova & McBride, 2021).

공공데이터는 사회, 경제적으로 다양한 이점을 제공할 수 있는데 정부 행정의 투명성과 책무성을 높이는 한편 새로운 비즈니스 기회 및 서비스 개발, 시민참여 등을 촉진하는데 기여할 수 있다(Open Government Data, 2015). 이러한 공공데이터의 이점을 활용하기 위해 각국 정부는 데이터 플랫폼을 구축하여 공공데이터 제공을 확대하고 있으며(Marjanovic & Cecez-Kecmanovic, 2020), 정보기술을 기반으로 한 공공데이터 포털은 공공데이터에 대한 접근성을 향상시킴으로써 정부 투명성 제고, 시민참여 확대 등에 기여하고 있다(Lnenicka & Nikiforova, 2021).

이와 같이 공공데이터의 활용 가치 및 중요성이 부각되면서 공공데이터에 대한 다양한 차원의 연구들이 진행되고 있다. 공공데이터 활용 관련 선행연구들은 공공데이터 속성(Kim & Kim, 2015; Park & Kim, 2015; Vetrò et al., 2016), 공공데이터 제공 플랫폼의 사용성(Cha, et al., 2017; Chen, et al., 2023; Benmohamed, et al., 2024; Weerakkody, et al., 2017), 공공데이터 기반 애플리케이션 개발 및 활용(Lee & Eom, 2023; Park & Oh, 2024), 공공데이터 정책 및 거버넌스(Kim, et al., 2015; Yoon, 2013; Yoon & Lee, 2017) 등 여러 차원에서 공공데이터 활용에 대해 조망하고 있다.

그러나, 확대되는 공공데이터 개방 추세와 달리 미흡

한 활용 수준은 문제점으로 지적되고 있다(Oh, et al., 2023). 2023년 기준 국내 공공데이터 포털에서 제공되는 공공데이터 개방 건수는 10년 전 대비 16.6배 증가했으나, Open API 형태로 제공되는 공공데이터 중 100명 이상이 활용을 신청한 데이터는 전체의 1/4 수준에 그치고 파일데이터의 경우에도 1/3 정도는 100회 이하 다운로드된 것으로 조사됐다(Kwon, 2023).

특히 정보기술과 인터넷 환경의 발달로 공공데이터 개방이 디지털 플랫폼을 중심으로 이루어지고 있어 데이터 플랫폼 활성화 필요성이 제기되고 있는 반면, 데이터 플랫폼 관점의 공공데이터 활용 관련 실증연구는 제한적인 상황이다. 또한, 데이터 플랫폼 활용 관련 선행연구들은 일부 사용자를 대상으로 한 설문조사 기반의 연구에 그치고 있는 실정이다(Cha, et al., 2017; Benmohamed, et al., 2024; Weerakkody, et al., 2017).

해의 주요 공공데이터 포털과 달리, 한국의 공공데이터 포털에서는 다양한 차원의 데이터 플랫폼 사용성 관련 지표를 확인할 수 있다. 이에 본 연구는 한국 공공데이터 포털에서의 실제 사용성 관련 데이터를 수집하여 국내외에서 실증한 바 없는 공공데이터 활용 영향요인에 대한 탐색적 연구를 시도해 보고자 한다.

II. 이론적 배경

1. 공공데이터와 데이터 플랫폼

2000년대 들어 정보의 디지털화와 인터넷의 확산은 정부 데이터를 민간에 공개하는 새로운 흐름을 만들어냈다. 2010년 버락 오바마 전 미국 대통령의 유엔(UN) 총회 연설을 계기로 정부 투명성 증진, 시민참여 활성화 등을 지향하는 열린정부(Open Government)의 국제적 협의체인 ‘열린정부파트너십(Open Government Partnership)’이 출범하였으며 2013년 5월 미 행정부는 ‘열린정부구상(Open Government Initiative)’을 통해 정부 정보에 대한 새로운 기본값으로 “개방적이

고 기계가 읽을 수 있는 데이터(Open and Machine-Readable Data)”를 설정하는 행정 명령을 시행했다(The White House, 2013). 이후 OECD가 ‘열린 정부를 위한 이사회 권고문(Recommendation of the Council on Open Government)’을 통해 쉽게 이용 가능한 공공 부문(Public Sector) 데이터 및 정보 제공을 회원국에 권고(OECD, 2017)하는 등 전세계적으로 공공데이터 개방이 확산되게 된다.

국내 공공데이터법에 따르면 공공데이터(Public Data)란 “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(공공데이터의 제공 및 이용 활성화에 관한 법률 제2조 제2호). 일반적으로 공공데이터는 정부기관이나 공공기관이 생성하거나 수집한 정보로 일반 대중들에게 자유롭게 제공되는 데이터로 정의되며(Attard, et al., 2015), 정보통신기술(ICT)의 발전과 더불어 공공데이터의 수집, 저장, 확산에 디지털 플랫폼이 적극적으로 활용되고 있다.

디지털 플랫폼 비즈니스 모델은 인터넷과 같은 디지털 인프라를 기반으로 제품 및 서비스의 공급자와 수요자를 비롯한 다양한 이해관계자들이 상호작용하며 가치를 만들어내는 비즈니스 형태라고 할 수 있다. 일반적인 디지털 플랫폼은 사용자 간 상호작용, 상품이나 서비스의 거래, 콘텐츠의 공유와 같은 기능을 제공하는데 중점을 두고 있는 반면 공공데이터 포털, 데이터 거래소 등을 포함한 데이터 플랫폼은 주로 데이터의 표준화, 보관, 공유 및 판매를 통해 정보의 접근성과 활용도를 높이기 위해 설계되었다는 점에서 구분된다.

데이터 플랫폼은 데이터를 생산, 수집, 가공, 분석하는 정보기술 인프라를 기반으로 데이터를 매개로 하여 다양한 참여자 간 협력을 통해 가치를 창출하는 체계를 의미하며(Ministry of Science and ICT, 2021) 데이터 기반의 가치를 창출하고 데이터를 생산, 분석, 거래하는 이용자들 간 상호작용이 이루어지는 공간으로 정의되기도 한다(Oh, et al., 2023). 특히 공공분야 데이

터 플랫폼은 “데이터를 입력·수집하고 처리·가공·분석 또는 유통·공유하여 새로운 공공가치 창출에 기여할 목적으로, 정부, 지자체, 공공 기관 등 공공분야 주체에 의해 구축·운영되는 통합된 디지털 환경, 시스템, 서비스, 인터페이스 등”으로 정의할 수 있는데(Oh, et al., 2023) 대표적인 공공데이터 플랫폼으로는 한국의 공공데이터포털(Data.go.kr)을 비롯하여 미국의 Data.gov, 영국의 Data.gov.uk 등이 있다.

2. 공공데이터 활용

선행연구들은 데이터셋, 데이터 플랫폼, 정부 정책 및 거버넌스 등 다양한 수준에서 공공데이터 활용 및 영향에 대해 연구하고 있다. 먼저 공공데이터셋 자체 속성 및 품질 관점에서 공공데이터 속성과 공공데이터 활용 간 관계에 대해 조망한 연구들이 있다. 일반적으로 데이터 품질(Data Quality)은 사용자가 사용하기에 적합한 수준, 사용자 편의에 부합하는 수준 등으로 정의된다(Park & Kim, 2015; Wang & Strong, 1996). 데이터 품질은 다차원적인 개념으로 이해되고 있으며 학자들 간에 일치된 견해는 부재하나 주로 언급되는 차원으로는 정확성, 완전성, 연관성, 가용성, 일관성, 적시성, 보안성 등이 있다(Quarati, 2023; Vetrò, et al., 2016; Wand & Wang, 1996).

다차원적인 데이터 품질에 대한 연구를 바탕으로 Park and Kim(2015)은 공공데이터가 갖춰야 할 품질 특성으로 공공성, 활용성, 신뢰성, 적합성 등을 도출한 바 있으며 Vetrò, et al.(2016)은 공공데이터 평가체계에 대한 연구에서 추적가능성, 최신성, 유효기간, 완전성, 규정준수, 이해가능성, 정확성 등을 제시하기도 하였다. Kim and Kim(2015)은 내재적, 접근적, 표현적 데이터 품질 요인을 중심으로 데이터 품질이 지각된 이용용이성과 지각된 유용성에 미치는 영향과 이를 매개로 공공데이터 이용 만족에 미치는 영향을 확인하였다.

한편, 데이터 플랫폼(포털) 관점에서 공공데이터셋의 자체 속성보다는 공공데이터를 제공하고 사용하는

시스템과 환경에 대한 분석을 중심으로 공공데이터 활용에 대해 조사한 연구들도 있다. Weerakkody, et al.(2017)은 공공데이터를 활용하는 시민 관점에서 공공데이터 플랫폼의 사용성(Usability)을 분석하고 플랫폼 개발에 있어 사용자 친화적 인터페이스 설계를 강조하고 있으며 Benmohamed, et al.(2024)은 데이터 플랫폼의 품질과 사용자 경험이 공공가치 창출에 중요한 요인임을 주장한 바 있다. Cha, et al.(2017)은 정보시스템 성공모델과 기술수용모델을 결합한 연구모형으로 공공데이터 포털 이용의도에 영향을 끼치는 요인을 검토하여 시스템 품질, 정보 품질, 정보 보안, 혁신성 등의 순으로 중요도를 확인하였다.

또한, 데이터 관리 정책, 거버넌스 체계, 법적·제도적 지원 등 정부 정책 및 거버넌스 차원에 초점을 둔 연구로 공공데이터 국가 연계 체계 구축(Yoon, 2013), 공공데이터 품질 관리를 위한 조직 성숙도 측정 모델(Kim, et al., 2015) 등을 제안하거나 데이터 개방 선진국 사례를 중심으로 공공데이터 거버넌스 체계(Yoon & Lee, 2017)에 대한 연구가 수행되었으며 G7 주요국 공공데이터 포털의 사용성을 비교 분석한 연구(Mutambik, et al., 2021) 등이 있다. 뿐만 아니라 공공데이터 수요 예측(Lee, 2020), 공공데이터 개방 및 활용 활성화 촉진요인과 저해요인 도출(Jeon & Kim, 2017) 등이 시도되기도 하였다.

III. 선행연구 및 가설설정

1. 공공데이터 플랫폼 평가기준

전 세계적으로 공공데이터 플랫폼 구축이 점차 확대되고 이를 중심으로 한 공공데이터 개방 사례는 증가하고 있으나 공공데이터 활용 수준은 여전히 미흡한 실정이다. 특히 선행연구들은 수집 가능한 자료의 제약으로 인해 실제 공공데이터 플랫폼 활용 관련 데이터를 이용해 실증분석하기 보다는 공공데이터 포털 디자인 및 평가체계를 개념적으로 제안(Francey & Mettler, 2022;

Máchová, et al., 2018; Rudmark, et al., 2024)하거나 제한된 사용자 샘플을 대상으로 공공데이터 활용 및 플랫폼 이용에 대한 설문조사에 의존하고 있다는 한계가 있다(Cha, et al., 2017; Chen, et al., 2023; Benmohamed, et al., 2024; Weerakkody, et al., 2017).

미국, 영국, 독일, 프랑스, 캐나다 등 해외 주요국의 대표적인 공공데이터 포털과 달리, 한국의 공공데이터 포털에서는 실제 공공데이터셋 다운로드 수뿐만 아니라 '좋아요'와 같은 사용자 평가 횟수 등 다양한 차원의 데이터 플랫폼 사용성 관련 지표를 확인할 수 있다. 이에 본 연구는 한국 공공데이터 포털에서의 실제 사용성 관련 데이터를 중심으로 공공데이터 활용 영향 요인에 대해 살펴보고자 한다.

Máchová, et al.(2018)은 공공데이터 플랫폼에서 데이터를 얼마나 쉽게 이용할 수 있는지를 의미하는 사용성(Usability)이 가장 일반적인 품질의 기준으로 활용될 수 있다고 강조하였는데 오픈데이터셋 사양(Specification), 오픈데이터셋 피드백(Feedback), 오픈데이터셋 요청(Request) 등을 중심으로 14개 세부 영역에서 오픈데이터 포털 평가체계를 주장한 바 있다. 이러한 공공데이터 플랫폼 평가체계는 국가별 공공데이터 포털에 대한 질적 비교 연구 등에 응용되어 활용되기도 하였다(Mutambik, et al., 2021; Nikiforova & McBride, 2021).

공공데이터 플랫폼 평가기준 중 오픈 데이터셋의 사양은 사용자가 데이터를 쉽게 발견하고 이해할 수 있도록 돕는 포털의 기능이며, 오픈 데이터셋 피드백은 데이터셋 사용 후 사용자가 의견을 제공하거나 데이터 품질을 개선할 수 있는 경로를 제공하는 것을 의미한다(Máchová, et al., 2018). 선행연구를 참고하여 본 연구에서는 한국 공공데이터 포털 사례를 중심으로 데이터셋 사양 관점에서 데이터셋 설명, 오픈포맷, 업데이트 수 등과 데이터셋 피드백 관점에서 사용자 평가가 실제 공공데이터셋 다운로드 수에 미치는 영향을 탐색적으로 살펴보고자 한다.

2. 데이터셋 설명

Máchová, et al.(2018)은 공공데이터 포털 평가체계 중 오픈데이터셋 사양과 관련한 세부적인 기준으로 데이터셋 설명, 데이터셋 발행자, 주제별 카테고리, 발행일 및 업데이트, 기계판독포맷, 오픈 라이선스, 시각화 및 분석도구 등을 제시하고 있다. 선행연구들은 데이터 플랫폼에서 고품질의 데이터셋, 충실한 메타데이터, 검색 기능 등이 사용자 만족에 중요하다고 강조한다(Francey & Mettler, 2022; Janssen, et al., 2012).

공공데이터셋에 대한 충실한 설명은 공공데이터셋의 사용성에 긍정적인 영향을 미칠 수 있다. 일반적으로 제품에 대한 설명은 사용자가 제품을 평가하고 결정을 내리는 데 도움이 되며, 자세한 설명은 제품에 대한 사용자의 불확실성을 제거하고 사용자가 제품에 대한 인상과 느낌을 형성하도록 돕는다(Khare & Rakesh, 2011; Mou, et al., 2020; Yang, et al., 2023).

Oh, et al.(2023)은 자료 및 정보 측면의 데이터플랫폼 활용 영향요인으로 데이터 양의 충분성, 데이터 다양성, 데이터 목적성, 데이터 표준화, 데이터 업데이트, 데이터 수요 반영 등을 제시하고 있다. 양적으로 충분한 데이터는 정보 품질을 확보하고 신규 정보 상품 및 서비스 개발에 기여할 수 있다는 점에서 중요하다.

국내 공공데이터 포털의 경우 개별 데이터셋에 대한 간략한 소개글과 더불어 데이터셋의 규모를 가늠할 수 있는 전체 행(Row) 수를 제시하여 이용자들의 데이터셋에 대한 이해를 돕고 있다. 상대적으로 긴 소개글과 큰 행 수는 데이터셋이 더 풍부하거나 상세할 것이라는 사용자 기대를 높여 데이터셋 활용에 긍정적인 영향을 미칠 수 있을 것으로 예상된다. 이에 본 연구는 데이터셋에 대한 텍스트 형태의 정성적 설명과 데이터셋 행수와 같은 정량적 설명이 공공데이터 활용에 긍정적인 영향을 미칠 수 있다는 점에서 다음과 같은 가설을 설정한다.

H1a: 정량적 공공데이터셋에 대한 설명은 공공데이

터 활용에 긍정적인 영향을 미칠 것이다.

H1b: 정성적 공공데이터셋에 대한 설명은 공공데이터 활용에 긍정적인 영향을 미칠 것이다.

3. 오픈포맷

종종 공공데이터와 오픈데이터(Open Data)가 혼용되어 사용되기도 하나 모든 공공데이터가 엄밀한 의미에서 오픈데이터의 요건을 충족하고 있는 것은 아니다. 글로벌 정보 공유 및 촉진을 도모하는 비영리조직인 열린지식재단(Open Knowledge Foundation)이 정의하는 ‘open’은 데이터나 콘텐츠가 누구에게나 자유롭게 접근, 사용, 수정, 공유될 수 있으며 그 목적에 제한이 없다는 원칙에 기초하고 있는데 이용성 및 접근성, 재사용과 재배포, 보편적 참여 등을 구체적 기준으로 제시하고 있다(Open Knowledge Foundation, 2015).

오픈데이터는 플랫폼에 독립적이고(Platform Independent), 기계가 읽을 수 있으며(Machine Readable), 정보를 재사용(Re-use)하는데 방해가 되는 제한없이 대중에게 제공될 수 있어야 한다는 측면에서 오픈라이선스, 접근성, 기계 가독성, 오픈포맷 등의 요건을 충족해야 한다(Open Knowledge Foundation, 2015; The White House, 2009). 오픈데이터는 디지털화되어 온라인에 표준화된 형식으로 제공되며 쉽게 검색하고 사용할 수 있는 데이터라는 측면에서 접근성 뿐만 아니라 활용성과 재사용 가능성에 중점을 두고 있다. 또한, 오픈데이터는 법적, 기술적인 측면의 개방성(Openness)을 요구하는데, 이는 데이터에 대한 자유로운 접근과 재사용을 위해 적절한 개방형 라이선스를 적용하거나 데이터를 공공 도메인에 배치하는 한편 데이터 사용에 기술적 장벽이 없어야 한다는 것을 의미한다(James, 2013).

국내 공공데이터 관리지침은 오픈포맷(Open Format)을 “기계 판독이 가능한 형태로서 비용 또는 그 밖의 사용에 제약 없이 최소 하나 이상의 소프트웨어로 처리할 수 있는 파일포맷”(공공데이터 관리지침, 제1장 제2조)

으로 규정하고 있으나 실제 공공데이터 포털 데이터셋의 기술적 개방도는 미흡한 사례를 다수 확인할 수 있다. 예를 들어, HWP, DOCX 등의 문서로 공공데이터를 제공하는 경우 전용 S/W를 사용하지 않고 데이터셋을 다루기는 어렵기 때문에 이를 오픈데이터라 평가할 수 없다.

특히 기술적인 관점에서 공공데이터의 개방성 정도는 공공데이터가 오픈포맷으로 제공되는지 여부로 평가해 볼 수 있는데 오픈포맷 준수 여부는 열린지식재단이 전세계 정부의 오픈데이터 정책을 평가하는 Global Open Data Index, 세계은행의 Open Data Readiness Assessment 등의 평가지표로 고려된다. 오픈포맷에 부합하는 공공데이터의 개방성 정도는 실제 공공데이터의 접근성, 활용성에 긍정적인 영향을 미칠 수 있을 것으로 기대되며 이에 다음과 같은 가설을 설정한다.

H2: 오픈포맷으로 제공되는 공공데이터는 공공데이터 활용에 긍정적인 영향을 미칠 것이다.

4. 업데이트

정보시스템의 특징적인 변화는 사용자의 태도에 영향을 미칠 수 있다(Mikalef, et al., 2020). 업데이트는 사용 행동에 영향을 미치며 사용자의 흥미를 회복하는 데 도움이 될 수 있다. 정보시스템의 새로운 버전이 출시되면 이는 사용자의 호기심을 자극하여 다시 다운로드하고 사용에 이르게 된다(Comino, et al., 2019; Gokgoz, et al., 2021). 대부분의 경우 업데이트는 정보 내용을 보완하고 사용자를 돕기 위한 것이므로 긍정적인 영향을 미칠 것으로 기대된다(Mikalef, et al., 2020).

업데이트 빈도는 사용자의 제품 평가에도 영향을 미칠 수 있는데(Fleischmann, et al., 2016) 클라우드펀딩 플랫폼 맥락에서 빈번한 업데이트를 통한 정보 제공이 펀딩 성과에 긍정적인 영향을 미친다는 선행연구도 있다(Mollick, 2014). 또한, Máchová, et al.(2018)은

데이터셋 업데이트 일시, 빈도 등을 공공데이터 포털 사용성 평가와 관련 주요 요소로 강조한다. 이에 본 연구는 데이터셋에 대한 업데이트가 데이터 활용에 긍정적인 영향을 미칠 것으로 예상하여 다음과 같은 가설을 설정한다.

H3: 공공데이터셋 업데이트 수는 공공데이터 활용에 긍정적인 영향을 미칠 것이다.

5. 사용자 평가

공공데이터 포털 평가 관련 메타데이터 및 데이터셋 자체에 중점을 두고 있는 연구에 비해 이해관계자들의 참여를 중요하게 고려하는 연구들도 있다. Máchová et al.(2018)은 공공데이터 포털 평가체계에서 오픈데이터셋 사양과 더불어 피드백(Feedback)을 고려하고 있는데 세부 기준으로는 문서화 및 튜토리얼, 포럼, 사용자 평점, 소셜미디어 공유 등을 제시하고 있다. Cha, et al.(2017)은 공공데이터 이용에서 정보 품질 및 사회적 영향의 중요성을 강조하고 있는데 사용자 평가에 포함된 정보 품질은 정보 출처의 신뢰성에 대한 사용자의 인식에 영향을 미쳐 공공데이터의 활용에 영향을 미칠 수 있다.

일반적으로 사용자는 다양한 정보 원천에서 제품 속성에 대한 정보와 제안을 얻고자 한다. 정보 제공자로서 사용자의 평가는 다른 사용자의 의사결정 과정에 영향을 미칠 수 있다. 사용자 평가는 시장 결과를 현저히 개선할 수도 있는데(Cui, et al., 2012; Park, et al., 2007) 대량의 사용자 평가는 제품에 대한 사용자의 관심을 더 잘 끌어들이고 사용자 신뢰를 높일 수 있으며(Zhu & Zhang, 2010), 제품의 선택 기회를 증가시킨다(Godes & Mayzlin, 2004; Zhou & Duan, 2016).

데이터 플랫폼 맥락에서 사용자는 공공데이터셋에 대한 종종 불완전한 제품 정보에 직면하게 되어 사용자 평가를 참고할 수 있는데 이는 데이터셋 다운로드 및 활용으로 이어질 수 있다. 또한 긍정적인 사용자 입소문은

간접적으로 더 많은 다운로드를 불러올 수도 있다(Lee, et al., 2008; Zhou & Duan, 2016). 이에 본 연구에서는 사용자 평가가 공공데이터 활용에 긍정적인 영향을 미칠 것으로 예상하고 다음과 같은 가설을 설정한다.

H4: 공공데이터셋에 대한 좋아요 수는 공공데이터 활용에 긍정적인 영향을 미칠 것이다.

IV. 연구방법

1. 데이터

국내 공공데이터는 공공데이터포털, 서울열린데이터

광장, K-ICT빅데이터센터, 기관별 포털 등 다양한 채널을 통해 제공되고 있다. 특히 2013년 공공데이터법 시행 이후 공공데이터 포털(www.data.go.kr)은 공공기관이 생성, 취득, 관리하고 있는 공공데이터를 파일데이터, 오픈 API, 표준데이터, 연계데이터 등 다양한 형태로 제공하고 있는 통합 창구 역할을 담당하고 있으며 국내 공공데이터 플랫폼 중 가장 많이 사용되고 있는 대표적인 데이터 플랫폼이다(Oh, et al., 2023).

공공데이터를 지속적, 반복적으로 사용하는 이용자는 오픈 API를 통해 정기적으로 데이터를 다운로드할 가능성이 높을 것으로 예상되는 반면 비정기적으로 접근하는 일반 사용자들은 파일데이터를 일회성으로 다운로드할 가능성이 높다. 2024년 7월 기준, 공공데이터



<그림 1> 파일데이터 목록

<Fig. 1> List of file data

포털내 전체 9만여 건의 데이터셋 중 오픈 API로 제공되는 데이터셋의 비중은 12.6%를 차지하고 있으며 파일데이터 비중은 74.9%로 다수를 차지하고 있다. 본 연구는 일반 사용자들의 데이터 플랫폼 활용 현황을 탐색적으로 살펴보기 위해 파일데이터 이용을 중심으로 데이터를 수집하여 실증분석을 시도한다.

공공데이터 포털내 '데이터 찾기' 메뉴에서 파일데이터 목록을 조회하면 <그림 1>과 같이 최근에 업데이트된 파일을 중심으로 파일데이터 목록을 조회할 수 있다.

개별 데이터셋별로 어떤 파일포맷(CSV, PDF, HWP 등)으로 제공되는지를 포함한 데이터명과 간단한 소개, 카테고리, 제공기관, 조회수, 다운로드수 등을 확인할 수 있다.

또한, 특정 파일데이터를 선택하면 <그림 2>와 같이 파일데이터에 대한 상세 내용을 조회할 수도 있다. 상세 내용에서는 데이터셋 소개글과 함께 좋아요, 싫어요와 같은 시민 평가 지표를 제공하고 있으며 하단의 파일데이터 정보에서는 DCAT 메타데이터뿐만 아니라 전체

CSV 기상청_날씨 이슈별 데이터(폭염)

바로가기

폭염이란 매우 심한 더위를 말하며, 폭염이 지속되면 온열질환 발생, 가축과 어패류의 집단 폐사, 야외근로자 사망 등 다양한 분야에서 피해가 발생합니다. 폭염 데이터셋은 폭염과 연관된 체감온도, 습도, 폭염 영향예보, 열대야 등 다양한 데이터를 하나로 묶어서 제공합니다.

👍 3

💬 0

👁️ 관심

파일데이터 정보 메타데이터 다운로드

데이터 개선요청

오류신고 및 문의

파일데이터명	기상청_날씨 이슈별 데이터(폭염)_20220819		
분류체계	과학기술 - 과학기술진흥	제공기관	기상청
관리부서명	국가기후데이터센터	관리부서 전화번호	042-481-7480
보유근거		수집방법	
업데이트 주기	수시 (자율 갱신)	자기 등록 예정일	
매체유형	텍스트	전체 행	1
확장자	CSV	키워드	날씨,폭염,기상현상
데이터 한계		다운로드(바로가기)	10260
등록일	2021-01-13	수정일	2022-08-19
제공형태	기관자재에서 다운로드(제공데이터URL기재)		
URL	https://data.kma.go.kr/data/weatherissue/silhtList.do?pgmNo=690		
설명	폭염이란 매우 심한 더위를 말하며, 폭염이 지속되면 온열질환 발생, 가축과 어패류의 집단 폐사, 야외근로자 사망 등 다양한 분야에서 피해가 발생합니다. 폭염 데이터셋은 폭염과 연관된 체감온도, 습도, 폭염 영향예보, 열대야 등 다양한 데이터를 하나로 묶어서 제공합니다.		
기타 유의사항	데이터는 5-9월에만 제공되며 매일 전일자료까지 업데이트되며 출단위 파일로 제공됩니다. 파일 다운로드는 로그인 후 이용하실 수 있습니다.		
비용부과유무	무료	비용부과기준 및 단위	건
이용허락범위	공공저작물_출처표시		

<그림 2> 파일데이터 상세 화면
<Fig. 2> Detailed view of file data

데이터셋 행 수, 담당기관, 제공/활용현황, 비용/제한사항 등을 포함한 세부적인 파일 데이터에 대한 설명을 추가적으로 제공한다.

2. 측정변수

본 연구는 공공데이터 포털의 활용과 영향 요인을 탐색적으로 살펴보기 위해 선행연구에서 활용한 공공데이터 플랫폼 평가기준을 참고하였다(Máková, et al., 2018; Mutambik, et al., 2021; Nikiforova & McBride, 2021). 실증분석을 위해 공공데이터 포털에서 제공하는 파일데이터 활용 현황을 중심으로 프로그래밍언어를 활용한 웹스크래이퍼(Web scraper)를 직접 제작하여 개별 데이터셋 단위로 자료를 수집했으며 주요 측정변수는 <표 1>과 같다.

종속변수인 공공데이터 활용 수준을 측정하기 위해 선행연구를 참고하여 실제 데이터셋 다운로드 수를 수집하였다(Ubaldi, 2013; Quarati, 2023). 주요 독립변수로 데이터셋의 정성적 설명을 측정하기 위해 데이터셋 소개글에 대한 문자 수를 활용하는 한편

정량적 설명으로 데이터셋의 전체 행 수를 측정하였다(Ehrlinger & WöB, 2022; Strong, et al., 1997; Wang & Strong, 1996). 오픈포맷은 제공되는 데이터셋 파일의 확장자 정보를 수집하고 선행연구의 오픈포맷 파일 확장자 분류기준을 적용하여 더미변수로 측정하였다(Máková, et al., 2018; Neumaier, et al., 2016; Wikipedia, n.d.). 또한, 사회적 평판 관련해서는 좋아요 수를 측정변수로 활용하였다(Máková, et al., 2018). 기타 통제변수로는 데이터셋 분류체계, 제공기관 유형, 데이터셋 게시기간 등을 분석에 활용하였다(Máková, et al., 2018; Mutambik, et al., 2021; Nikiforova & McBride, 2021).

3. 모형설정

가설검증을 위해 활용한 회귀분석 모형은 식 (1)과 같다. 다운로드 수, 소개글 문자 수, 전체 행 수, 좋아요 수, 업데이트 수, 게시기간 등의 연속형 변수는 로그변환하여 추정에 활용하였다.

<표 1> 측정변수
<Table 1> Measurement

Category	Variable	Description
Dependent variable	Downloads	Number of dataset downloads
Independent variable	Description	Number of characters in dataset description
	NumRows	Number of rows in the dataset
	OpenFormat	Open format = 1; Non-open format = 0 * open format file extensions: bmp, csv, dbf, epub, gif, gltf, gpkg, gpx, html, json, jpeg, jpg, odt, pdf, png, rdf, rss, tiff, ttl, txt, xml, zip, flac, webm, svg, ogg, wav, mp3, etc.
	Update	Number of updates
	Like	Number of likes
Control variable	Duration	Posting duration (days)
	Category	Dataset category
	Institution	Type of dataset publishing institution

$$\begin{aligned}
 \text{LnDownload} = & \beta_0 + \beta_1 * \text{LnDescCnt} \\
 & + \beta_2 * \text{LnNumRows} \\
 & + \beta_3 * \text{OpenFormat} \\
 & + \beta_4 * \text{LnLike} \\
 & + \beta_5 * \text{LnUpdates} \\
 & + \beta_6 * \text{LnDuration} \\
 & + \beta_7 * \text{Category} \\
 & + \beta_8 * \text{Institution} + \varepsilon
 \end{aligned}
 \tag{1}$$

고리별로는 공공행정(9,315건, 비중 13.9%), 문화관광(8,250건, 12.3%), 산업고용(6,459건, 9.7%) 등이 상위 를 차지한 것으로 나타났다. 제공기관별로는 자치행정 기관이 35,903건(비중 53.7%)으로 데이터를 가장 많이 공개하고 있는 가운데 공공기관(20,661건, 30.9%), 국가행정기관(9,790건, 14.6%)이 뒤를 이었다.

공공데이터셋의 평균 다운로드 건수는 639건으로 가장 많은 것은 1,158,187건에 달했으나 한번도 다운로드되지 않은 데이터셋도 있는 것으로 나타났다. 소개글 문자 수와 데이터셋 행 수의 평균은 각각 92.5, 61,455.6 등으로 집계됐다. 데이터셋에 대한 가장 많은 좋아요 수는 48건이었으나 평균적으로는 좋아요 수가 없는 것이 다수인 것으로 나타났다. 한편, 데이터셋 평균 업데이트 수는 2.2건이었으며 게시기간은 평균 913일로 나타났다.

분석에 활용된 주요 변수 간 상관관계는 <표 3>과 같다. 다운로드 건수(LnDownload)는 데이터셋 행 수

V. 분석결과

1. 기술통계

본 연구에서 활용된 공공데이터셋은 66,871건으로 기술통계량은 <표 2>와 같다. 오픈포맷으로 제공되는 공공데이터셋 수는 58,918건으로 전체 데이터셋의 88.1%의 비중을 차지하고 있는 것으로 나타났다. 카테고리

<표 2> 기술통계량
<Table 2> Descriptive statistics

Variable	Obs	Mean	Std. dev.	Min	Max
Downloads	66,871	639.1	4,700.8	0	1,158,187
Description	66,871	92.5	50.9	5	1,000
NumRows	66,871	61,455.6	2,286,487.0	0	244,000,000
Update	66,871	2.2	6.6	1	1,251
Like	66,871	0.0	0.4	0	48
Duration	66,871	913.2	855.8	7	4,147
OpenFormat	Open format: 58,918 datasets (88.1%), Non-open format: 7,953 (11.9%)				
Category	Public Administration: 9,315 datasets (13.9%), Culture and Tourism: 8,250 (12.3%), Industry and Employment: 6,459 (9.7%), Transportation and Logistics: 5,566 (8.3%), Environment and Weather: 5,463 (8.2%), National Territory Management: 4,687 (7.0%), Financial and Taxation: 4,297 (6.4%), Agriculture and Fisheries: 3,999 (6.0%), Social Welfare: 3,778 (5.7%), Health and Medicine: 3,633 (5.4%), Disaster Safety: 3,562 (5.3%), Education: 3,216 (4.8%), Science and Technology: 1,726 (2.6%), Food Safety: 1,507 (2.3%), Unification and Foreign Affairs: 970 (1.5%), Legislation: 443 (0.7%)				
Institution	Local Administrative Agencies: 35,903 datasets (53.7%), Public Institutions: 20,661 (30.9%), National Administrative Agencies: 9,790 (14.6%), Educational Administrative Agencies: 354 (0.5%), Committees: 119 (0.2%), Constitutional Institutions: 28 (0.0%), Educational Institutions: 15 (0.0%), Legislative Institutions: 1 (0.0%)				

〈표 3〉 주요 변수 간 상관관계
 〈Table 3〉 Correlations of key variables

	[1]	[2]	[3]	[4]	[5]	[6]
[1] LnDownload	1.00					
[2] LnDescCnt	-0.05***	1.00				
[3] LnNumRows	0.05***	-0.03***	1.00			
[4] LnUpdates	0.45***	-0.00	0.10***	1.00		
[5] LnLike	0.12***	0.03***	0.04***	0.06***	1.00	
[6] LnDuration	0.12***	-0.08***	-0.14***	-0.51***	0.00	1.00

〈표 4〉 실증분석 결과
 〈Table 4〉 Analysis results

	Model(1)	Model(2)	Model(3)	Model(4)	Model(5)	Model(6)
LnNumRows	0.035*** (0.002)	0.035*** (0.002)				0.018*** (0.002)
LnDescCnt	-0.139*** (0.015)	-0.775*** (0.124)				-0.838** (0.100)
LnDescCntSq		0.071*** (0.014)				0.090*** (0.011)
OpenFormat			0.580*** (0.019)			0.253*** (0.015)
LnUpdates				1.597*** (0.008)		1.564*** (0.008)
LnLike					1.566*** (0.052)	0.930*** (0.042)
LnDuration	0.172*** (0.005)	0.168*** (0.005)	0.186*** (0.005)	0.603*** (0.005)	0.165*** (0.005)	0.603*** (0.005)
Category	Yes	Yes	Yes	Yes	Yes	Yes
Institution	Yes	Yes	Yes	Yes	Yes	Yes
Constant	5.074*** (0.079)	6.504*** (0.288)	3.986*** (0.041)	1.052*** (0.035)	4.597*** (0.036)	2.679*** (0.232)
N	66,871	66,871	66,871	66,871	66,871	66,871
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2	0.0332	0.0336	0.0419	0.3680	0.0411	0.3779
Adj R2	0.0328	0.0332	0.0415	0.3678	0.0408	0.3776

note: Standard errors are shown in parentheses below the estimates.

* p<0.1, ** p<0.05, *** p<0.01

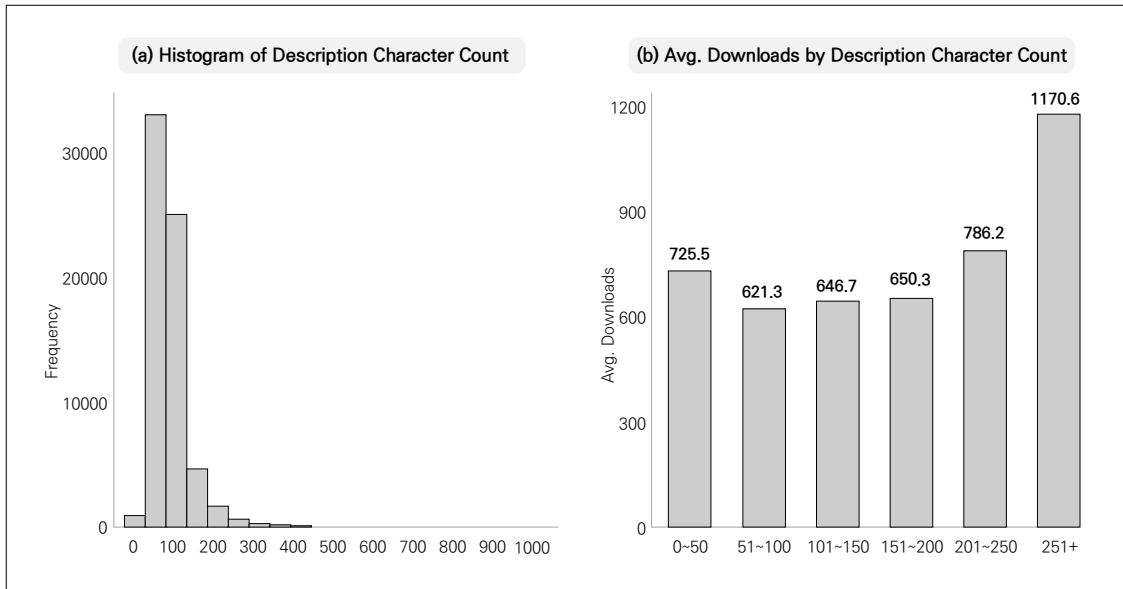
(LnNumRows), 업데이트 수(LnUpdates), 좋아요 수(LnLike), 게시기간(LnDuration) 등과 양의 상관관계를 보이고 있는 반면 소개글 문자 수(LnDescCnt)와는 음의 상관관계가 있는 것으로 나타났다. VIF값 측정 결과, 주요 변수들의 VIF 값은 1.00~1.71 사이에 분포하고 평균 VIF값이 1.27로 측정되어 모형에 활용된 변수 간 다중공선성 이슈는 없는 것으로 확인되었다.

2. 가설검증

가설검증 결과는 <표 4>와 같다. 모형 (1)~(5)는 통계 변수를 고려하여 개별 가설을 검증한 결과이며 모형 (6)은 전체 가설에 대한 검증결과이다. 먼저 모형 (1)에서 데이터셋 행 수는 다운로드 수에 긍정적인 영향(coef. =.035, $p < .01$)을 미치는 것으로 분석되어 가설 H1a를 지지하는 것으로 나타났다. 한편, 소개글 문자 수가 데이터셋 다운로드에 미치는 영향과 관련 초기 분석 결과에서는 예상과 달리 부(-)의 관계가 나타났다(coef.

= -.139, $p < .01$). 이에 모형 (2)에서와 같이 소개글 문자 수에 대한 이차항을 추가적으로 고려하여 회귀분석한 결과, 이차항의 회귀계수는 +0.071($p < .01$)로 유의하게 추정되었다. 데이터셋에 대한 소개글이 적은 구간에서는 다운로드 건수가 유의미하게 증가하지는 않지만 소개글이 어느 정도 이상으로 충분히 많아지면 다운로드 건수가 증가하는 패턴을 갖는다고 해석되어 가설 H1b는 부분적으로 지지되고 있다.

이와 같은 패턴은 <그림 3>과 같이 소개글 수에 대한 분포를 시각화한 자료에서도 확인된다. <그림 3(a)>는 공공데이터셋의 전반적인 소개글 수 히스토그램으로 많은 데이터셋들의 소개글 수가 150자 이내에 집중되고 있는 것으로 집계된다. 소개글 수에 대한 회귀계수 추정에서 해당 구간의 자료들이 큰 영향을 미칠 수 있음을 가늠해 볼 수 있다. 한편, <그림 3(b)>는 소개글 수 구간별 평균 데이터셋 다운로드 건수를 보여준다. 데이터셋 소개글 수가 100자 이내인 그룹에서도 다운로드 건수는 발생하고 있으나 소개글 수가 증가하더라도 데이



〈그림 3〉 데이터셋 소개글 수 분포
(Fig. 3) Distributions of description character count

터셋 다운로드 건수가 증가하는 패턴을 확인하기는 어렵다. 데이터셋 다운로드에서 소개글 이외에 다른 요인에 의해 영향을 받았을 가능성이 있다. 반면, 소개글 수가 100자를 초과한 그룹에서는 소개글 수가 증가할수록 평균 다운로드 건수가 점차 증가하는 양상을 보이며 특히 200자를 초과한 설명에서 그 효과가 더욱 뚜렷하게 나타난다. 즉, 질적 측면에서 일정 수준 이상의 충분한 소개글이 데이터셋 다운로드와 긍정적인 관계를 갖는다고 해석해 볼 수 있다.

(표 4)의 모형 (3)에서 오픈포맷은 다운로드 수에 긍정적인 영향(coef. = .580, $p < .01$)을 미치는 것으로 나타났다. 오픈포맷으로 제공되는 공공데이터셋이 데이터 활용에 긍정적인 영향을 미친다는 가설 H2를 지지하는 것으로 나타났다. 개방형 포맷을 활용한 공공데이터 제공이 공공데이터 이용 활성화에 중요함을 시사한다. 모형 (4)에서 업데이트 수의 회귀계수 또한 양(+)의 값(coef. = 1.597, $p < .01$)으로 가설 H3를 지지하는 것으로 나타났다. 데이터셋 업데이트 수는 공공데이터 활용에 긍정적인 영향을 미친다는 것으로 해석된다. 모형 (5)에서 좋아요 수의 회귀계수는 양(+)의 값(coef. = 1.566, $p < .01$)으로 나타나 가설 H4를 지지하고 있다. 이는 '좋아요'와 같은 사용자 평가가 공공데이터 활용에 긍정적인 영향을 미친다는 것으로 해석할 수 있다.

모형 (6)은 통제변수와 주요 독립변수를 종합적으로 고려한 모형으로 앞서 모형 (1)~(5)에서 확인한 내용과 일관된 분석결과를 보여주고 있다. 이상의 전체적인 회귀분석 결과에 따르면 가설 H1b가 부분적으로 지지된 가운데 이외의 가설들은 지지되는 것으로 나타났다.

VI. 결론 및 시사점

본 연구는 공공데이터 포털에서 제공하는 파일데이터를 중심으로 국내외에서 시도된 바 없는 공공데이터 활용 영향요인에 대한 실증분석을 수행하였다. 분석결

과에 따르면 정량적인 데이터셋에 대한 설명과 오픈포맷 형태의 데이터 제공, 사용자 평가, 데이터셋 업데이트 등이 공공데이터 활용에 긍정적인 영향을 미치는 것을 확인하였다. 정성적인 데이터셋에 대한 설명은 설명 문자수가 적을 때는 다운로드 건수가 유의미하게 증가하지는 않지만 설명 문자수가 충분히 커질수록 다운로드 수가 증가하는 것으로 나타났다. 본 연구는 데이터 플랫폼의 활용성을 높이는 주요 요인들을 실증적으로 분석함으로써 공공데이터 정책 및 플랫폼 운영 개선에 다음과 같은 시사점을 제공해 준다.

먼저 데이터셋에 대한 너무 짧은 설명은 데이터셋 이용에 긍정적인 영향을 미치지 어려우므로 데이터셋 소개글은 충분히 상세하게 제공하는 것이 필요하다. 본 연구에서 살펴본 국내 공공데이터 포털 사례의 경우, 데이터셋 설명의 길이가 일정 수준 이상으로 길어질수록 평균 다운로드 건수는 점진적으로 증가하는 경향을 보이며 특히 200자를 초과한 설명에서 그 효과가 더욱 뚜렷하게 나타난다. 즉, 설명의 상세화가 다운로드 건수에 긍정적 영향을 미칠 수 있음을 보여준다.

둘째, 공공데이터 제공 시 오픈포맷을 사용하는 것이 중요하다. 비개방형 포맷으로 제공된 데이터는 사용자들이 분석과 가공에 어려움을 겪어 데이터 활용을 저해할 수 있다. 따라서 공공기관은 가능한 많은 데이터를 오픈포맷으로 제공해야 하며, 이를 통해 공공데이터의 접근성과 활용성을 동시에 높이는 것이 필요하다. 선행 연구는 PDF나 이미지 형태의 데이터 개방은 공개 이상의 의미를 갖기 어려우며 이용 활성화를 위해서는 최소한 RDF 변환이 가능한 수준으로 공공데이터가 제공되어야 한다고 강조하고 있다(Yi & Nam, 2014).

셋째, 데이터셋 업데이트 빈도도 공공데이터 활용에 중요한 요소로 작용한다. 데이터셋의 최신성은 사용자들이 신뢰할 수 있는 데이터를 제공받는 데 중요한 역할을 하므로 데이터를 관리하는 기관은 정기적으로 데이터를 업데이트하고 이를 사용자에게 알리는 노력이 필요하다.

마지막으로 데이터셋에 대한 사용자 평가, 특히 '좋

아요'와 같은 긍정적인 피드백은 데이터 활용을 촉진하는 데 중요한 역할을 한다. 이는 사용자들이 데이터셋의 품질과 유용성을 쉽게 판단할 수 있게 해주며, 더 많은 사용자들에게 신뢰를 주어 데이터 활용을 유도할 수 있다. 그러나, 국내 공공데이터 플랫폼 활용 현황을 살펴보면 데이터셋 사용자 평가는 아직 활성화되지 못한 것으로 나타나고 있다. 공공데이터 플랫폼 운영자는 사용자 평가 시스템을 적극 활용하고 이를 데이터셋 품질 개선에도 반영하는 것이 필요하다.

한편, 본 연구는 한국의 공공데이터 포털 사례만을 대상으로 하여 연구결과의 일반화에서 다소 제한적인 측면이 있으며 웹 스크레이핑 방식을 활용한 데이터 수집으로 데이터 품질의 정확성, 수집시점에 따른 데이터 편향 등의 문제가 있을 수 있다. 또한, 개별 데이터셋의 내용적 특성을 보다 상세히 고려하지 못한 것도 본 연구의 한계라 할 수 있다. 후속 연구에서는 다양한 국가의 공공데이터 플랫폼 사례를 추가하는 한편 파일데이터 이외에 오픈 API 형태의 데이터와 내용 분석을 포함하여 공공데이터 플랫폼 활용에 미치는 영향 요인을 보다 종합적으로 검토할 필요가 있다.

■ References

- Attard, J., Orlandi, F., Scerri, S. & Auer, S. (2015). "A systematic review of open government data initiatives." *Government Information Quarterly*, 32(4), 399-418.
- Benmohamed, N., Shen, J. & Vlahu-Gjorgievska, E. (2024). "Public value creation through the use of open government data in Australian public sector: A quantitative study from employees' perspective." *Government Information Quarterly*, 41(2), 101930.
- Cha, Y., Choi, S. & Han, K. (2017). "An Empirical Study on the Influence on Public Data Usage in Private Business Sectors." *Journal of Digital Convergence*, 15(6), 9-17.
- {차영일 · 최성규 · 한경석 (2017). 민간의 공공데이터 활용을 위한 이용의도에 미치는 영향에 관한 실증적 연구. <디지털융복합연구>, 15권 6호, 9-17.}
- Chen, M., Cao, Y. & Liang, Y. (2023). "Determinants of open government data usage: Integrating trust theory and social cognitive theory." *Government Information Quarterly*, 40(4), 101857.
- Comino, S., Manenti, F. M. & Mariuzzo, F. (2019). "Updates management in mobile applications: iTunes versus Google Play." *Journal of Economics & Management Strategy*, 28(3), 392-419.
- Cui, G., Lui, H. & Guo, X. (2012). "The effect of online consumer reviews on new product sales." *International Journal of Electronic Commerce*, 17(1), 39-58.
- Ehrlinger, L. & Wöb, W. (2022). "A Survey of Data Quality Measurement and Monitoring Tools." *Frontiers in Big Data*, 5, 850611.
- Fleischmann, M., Amirpur, M., Grupp, T., Benlian, A. & Hess, T. (2016). "The Role of Software Updates in Information Systems Continuance—An Experimental Study from a User Perspective." *Decision Support Systems*, 83, 83-96.
- Francey, A. & Mettler, T. (2022). *Satisfaction with Open Government Data Portals: A User-Centric Configurational Perspective. Forty-Third International Conference on Information Systems*. ICIS 2022 Proceedings, 2.
- Godes, D. & Mayzlin, D. (2004). "Using online conversations to study word of mouth communication." *Marketing Science*, 23(4), 545-560.
- Gokgoz, Z. A., Ataman, M. B. & van Bruggen, G. H. (2021). "There's an app for that! Understanding the drivers of mobile application downloads." *Journal of Business Research*, 123, 423-437.
- Janssen, M., Charalabidis, Y. & Zuiderwijk, A. (2012). "Benefits, Adoption Barriers and Myths of Open Data and Open Government." *Information Systems Management*, 29(4), 258-268.
- Jeon, B. & Kim, H. (2017). "An Exploratory Study on the Sharing and Application of Public Open Big Data." *Informatization Policy*, 24(3), 27-41.
- {전병진 · 김희웅 (2017). 공공 빅데이터 개방 및 활용 활성화 방안에 대한 연구. <정보화정책>, 24권 3호, 27-41.}

- Khare, A. & Rakesh, S., (2011). "Antecedents of online shopping behavior in India: an examination." *Journal of Internet Commerce*, 10(4), 227-244.
- Kim, H. & Kim, G. (2015). "A Study on Public Data Quality Factors Affecting the Confidence of the Public Data Open Policy." *Journal of Information Technology Services*, 14(1), 53-68.
- {김현철·김광용 (2015), 공공데이터 품질 요인이 공공데이터 개방정책의 신뢰에 미치는 영향에 관한 연구. <한국 IT 서비스학회지>. 14권 1호, 53-68.}
- Kim, S., Lee, C., Chung, S., Kim, H. & Lee, C. (2015). "An Organizational Maturity Assessment Model for Public Data Quality Management." *Informatization Policy*, 22(1), 28-46.
- {김선호·이창수·정승호·김학철·이창수 (2015), 공공데이터 품질관리를 위한 조직 성숙도 평가 모델. <정보화정책>, 22권 1호, 28-46.}
- Kubler, S., Robert, J., Neumaier, S., Umbrich, J. & Le Traon, Y. (2018). "Comparison of metadata quality in open data portals using the Analytic Hierarchy Process." *Government Information Quarterly*, 35(1), 13-29.
- Kwon, H. (2023). "The 'openness' of public data continues to increase and the 'utilization' is insufficient." <https://www.etnews.com/20230112000222>. (Retrieved on October 1, 2024).
- {권혜미 (2023). "공공데이터 '개방' 지속 증가... '활용'은 미흡." <https://www.etnews.com/20230112000222>. (검색일 : 2024.10.1)}
- James, L. (2013). "Defining Open Data." *Open Knowledge*. <https://blog.okfn.org/2013/10/03/defining-open-data/>. (Retrieved on October 1, 2024).
- Lee, J. (2020). "Forecasting Open Government Data Demand Using Keyword Network Analysis." *Informatization Policy*, 27(4), 24-46.
- {이재원 (2020), 키워드 네트워크 분석을 이용한 공공데이터 수요 예측. <정보화정책>, 27권 4호, 24-46.}
- Lee, J., Park, D. H. & Han, I. (2008). "The effect of negative online consumer reviews on product attitude: An information processing view." *Electronic commerce research and applications*, 7(3), 341-352.
- Lee, S. & Eom, S. (2023). "Analysis of Open Government Data Utilization Performance and Determinants: A focus on mobile apps using open government data." *Korean Journal of Public Administration*, 61(1), 41-76.
- {이수경·엄석진 (2023), 공공데이터 활용 성과와 결정요인 분석: 공공데이터를 활용한 모바일 앱을 중심으로. <행정논총>, 61권 1호, 41-76.}
- Lnenicka, M. & Nikiforova, A. (2021). "Transparency -by-design: What is the role of open data portals?" *Telematics and Informatics*, 61. 101605.
- Máková, R., Hub, M. & Lnenicka, M. (2018). "Usability evaluation of open data portals: Evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective." *Aslib Journal of Information Management*, 70(3), 252-268.
- Marjanovic, O. & Cecez-Kecmanovic, D. (2020). "Open government data platforms-A complex adaptive sociomaterial systems perspective." *Information and Organization*, 30(4), 100323.
- Mikalef, P., Krogstie, J., Pappas, I. O. & Pavlou, P. (2020). "Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities." *Information & Management*, 57(2), 103169.
- Ministry of Science and ICT (2021). "Development strategy of basic data platform for civil and government cooperation." <https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=112&pageIndex=&bbsSeqNo=94&nntSeqNo=3180352&searchOpt=ALL&searchTxt=>. (Retrieved on June 11, 2021).
- {과학기술정보통신부 (2021). "민·관 협력 기반 데이터 플랫폼 발전전략." <https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=112&pageIndex=&bbsSeqNo=94&nntSeqNo=3180352&searchOpt=ALL&searchTxt=>. (검색일 : 2021.06.11)}
- Mollick, E. (2014). "The Dynamics of Crowdfunding: An Exploratory Study." *Journal of Business Venturing*, 29(1), 1-16.
- Mou, J., Zhu, W. & Benyoucef, M. (2020). "Impact of product description and involvement on purchase intention in cross-border e-commerce." *Industrial*

- Management & Data Systems*, 120(3), 567-586.
- Mutambik, I., Almuqrin, A., Lee, J., Zhang, J. Z., Alomran, A., Omar, T., Floos, A. & Homadi, A. (2021). "Usability of the G7 Open Government Data Portals and Lessons Learned." *Sustainability*, 13(24), 13740.
- Neumaier, S., Umbrich, J. & Polleres, A. (2016). "Automated Quality Assessment of Metadata across Open Data Portals." *Journal of Data and Information Quality*, 8(1), 1-29.
- Nikiforova, A. & McBride, K. (2021). "Open government data portal usability: A user-centred usability analysis of 41 open government data portals." *Telematics and Informatics*, 58, 101539.
- OECD, (2017). "Recommendation of the Council on Open Government." <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0438>. (Retrieved on December 14, 2021).
- Oh, Y., Cha, N. & Kim, J. (2023). *Enhancing the utilization of data platforms in the public sector*. KIPA, 2023-12-31, 1-254.
- {오윤경 · 차남준 · 김준형 (2023). <공공분야 데이터플랫폼 활용성 제고방안 연구>. 한국행정연구원, 2023-12-31, 1-254.}
- Open government data. (2015). "In OECD, Government at a Glance 2015, 150-151. OECD." https://doi.org/10.1787/gov_glance-2015-48-en. (Retrieved on July 06, 2015).
- Open Knowledge Foundation (2015). "The Open Definition." <https://opendefinition.org/>. (Retrieved on 2015).
- Park, D. Lee, J. & Han, I. (2007). "The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement." *International Journal of Electronic Commerce*, 11(4), 125-148.
- Park, G. & Kim, C. (2015). "Quality Characteristics of Public Open Data." *Journal of Digital Convergence*, 13(10), 135-146.
- {박고은 · 김창재 (2015). 공공개방데이터 품질 특성에 관한 연구. <Journal of Digital Convergence>, 13권 10호, 135-146.}
- Park, T. & Oh, S. (2024). "A Study on the Intention to Use of Public Application : Focused on Publicness and Technology Readiness Acceptance Model(TRAM)." *Information Systems Review*, 26(2), 95-121.
- {박태환 · 오세환 (2024). 공공앱 사용의도에 관한 연구: 공공성과 기술준비수용모델을 중심으로. <경영정보학연구>, 26권 2호, 95-121.}
- Quarati, A. (2023). "Open Government Data: Usage trends and metadata quality." *Journal of Information Science*, 49(4), 887-910.
- Rudmark, D., Lindgren, R. & Schultze, U. (2024). Open data platforms: Design principles for embracing outlaw innovators. *The Journal of Strategic Information Systems*, 33(3), 101850.
- Strong, D. M., Lee, Y. W. & Wang, R. Y. (1997). "Data quality in context." *Commun. ACM*, 40(5), 103-110.
- The White House, (2009). "Open Government Directive." <https://obamawhitehouse.archives.gov/open/documents/open-government-directive>. (Retrieved on December 8, 2009).
- The White House, (2013). "Open Government Initiative." <https://obamawhitehouse.archives.gov/open> (Retrieved on May 9, 2013).
- Ubaldi, B. (2013). "Open government data: Towards empirical analysis of open government data initiatives." *OECD Working Papers on Public Governance*, 22.
- Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R. & Morando, F. (2016). "Open data quality measurement framework: Definition and application to Open Government Data." *Government Information Quarterly*, 33(2), 325-337.
- Wand, Y. & Wang, R. Y. (1996). "Anchoring data quality dimensions in ontological foundations." *Communications of the ACM*, 39(11), 86-95.
- Wang, R. Y. & Strong, D. M. (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers." *Journal of Management Information Systems*, 12(4), 5-33.
- Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U. & Dwivedi, Y. K. (2017). "Open data and its usability: An empirical view from the Citizen's

- perspective.” *Information Systems Frontiers*, 19(2), 285-300.
- Wikipedia. (n.d.), “List of open file formats.” https://en.wikipedia.org/wiki/List_of_open_file_formats (Retrieved on October 1, 2024).
- Yang, X., Ma, Y. & Han, J. (2023). “Product information the more the better? A moderated mediation model to explain consumer purchasing intention in cross-border e-commerce.” *Asia Pacific Journal of Marketing and Logistics*, 35(11), 2601-2620.
- Yi, H. & Nam, Y. (2014). “A Study on Revitalizing the Use of Korean Public Data: Focused on Linked Open Data Strategy.” *Korean Society for Information Management*, 31(4), 249-266.
- {이현정·남영준 (2014). 우리나라 공공데이터의 이용활성화 방안에 관한 연구: 링크드 오픈 데이터화 전략을 중심으로. <정보관리학회지>, 31권 4호, 249-266.}
- Yoon, K. & Lee, G. (2017). “Exploring the Governance Enhancing Opening the Public Data: Using the Case of ADRN in the United Kingdom.” *Institute of Public Policy and Administration*, 31(1), 1-27.
- {윤광석·이건 (2017). 공공데이터 활용 행정 촉진을 위한 거버넌스 모색: 영국의 ADRN 사례를 중심으로. <국가정책연구>, 31권 1호, 1-27.}
- Yoon, S. (2013). “A Study on National Linking System Implementation based on Linked Data for Public Data.” *Korea Society for Information Management*, 30(1), 259-284.
- {윤소영 (2013). 공공데이터 활용을 위한 링크드 데이터 국가연계체계 구축에 관한 연구. <정보관리학회지>, 30권 1호, 259-284.}
- Zhou, W. & Duan, W. (2016). “Do Professional Reviews Affect Online User Choices Through User Reviews? An Empirical Study.” *Journal of Management Information Systems*, 33(1), 202-228.
- Zhu, F. & Zhang, X. (2010). “Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics.” *Journal of marketing*, 74(2), 133-148.