

기반시설 마스터데이터 표준요소 구축에 관한 연구*

- 기반시설 표준데이터를 중심으로 -

A Study on the Establishment of Standard Elements of Infrastructure Master Data: Focused on Infrastructure Standard Dataset

손 혜 인 (Hyein Sohn)**

남 영 준 (Young Joon Nam)***

초 록

마스터데이터는 기관 내부의 광범위한 이용을 목적으로 구축되며, 주로 기업에서 많이 활용되고 있는 분야이다. 이 연구는 국가의 공공기관에서 활용할 수 있는 기반시설에 관한 마스터데이터 구축을 목적으로 연구를 진행하였다. 이를 위해 공공데이터포털에서 제공하는 표준데이터세트에 기반을 두고 해당 데이터세트의 개별 속성을 분석하였다. 이 중 마스터데이터의 특성에 맞는 표준요소를 추출하였고, 최종적으로 종합한 표준요소를 국가에서 활용하고 있는 표준화 체계를 통하여 검증하였다.

ABSTRACT

The Master Data is constructed for the wide use within the institution, and it is mainly used in the enterprise. In this research, we have conducted research for the purpose of building master data on infrastructure that can be used by public institutions in the country. To do this, we analyzed individual attributes of the standard data set provided by the public data portal. Among these, we extracted standard elements that match the characteristics of the Master Data. Finally, the standardized elements are verified through the standardization system that is utilized in the country.

키워드: 마스터데이터, 기반시설, 표준데이터, 데이터세트, 마스터데이터 구축, 공공데이터

Master Data, Infrastructure, Standard Dataset, Dataset, Master Data Establishment, Public Data

* 본 논문은 2017년도 석사학위논문을 토대로 수정·정리한 내용임.

** 중앙대학교 문헌정보학과 대학원 기록관리전공 석사(shyein92@cau.ac.kr) (제1저자)

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

논문접수일자: 2017년 11월 20일 논문심사일자: 2017년 12월 20일 게재확정일자: 2017년 12월 23일
한국비블리아학회지, 28(4): 35-55, 2017. [http://dx.doi.org/10.14699/kbiblia.2017.28.4.035]

1. 서론

1.1 연구의 필요성과 목적

대한민국 정부는 2001년 이후 행정기관의 생산성, 투명성 및 민주성을 높여 지식정보화시대에 국민 삶의 질을 향상하기 위한 목적으로 국가 차원에서 전자정부를 추진하였다. 이에 따라 전자정부 하의 모든 정부 기관과 공공기관은 행정정보를 생성함에 있어 기관별 특성에 따라 기록물관리 기준과 데이터 관리체계를 구축하게 되었다. 개별기관별로 데이터를 관리하는 현행 공공기관 데이터관리시스템은 전체 기관 측면에서 보았을 때 기록물과 행정정보관리의 일관성과 신뢰성, 데이터 통합관리 및 활용적 측면에서의 비용 및 효율성 문제를 갖는다. 이러한 문제점에 대한 대안으로 통합적인 데이터 관리체계가 필요하다.

이에 본 연구에서는 다양한 데이터 통합 방식 중 정보에 기반을 둔 통합 방식인 마스터데이터(Master Data)를 활용하여 공공시설에서 활용할 수 있는 기반시설 마스터데이터 구축 방안을 도출하고자 한다. 마스터데이터를 활용한 데이터 통합 방식은 현존하는 데이터뿐만 아니라 미래의 이용 가능성에 중점을 두고 데이터를 통합하는 방식이다. 기본적으로 마스터데이터는 광범위한 이용에 목적을 두고 구축되기 때문에 메타데이터나 디렉토리와 같은 기존의 데이터 통합 방식보다 구축된 데이터의 활용성이 높다. 따라서 본 연구의 목적은 정부와 공공기관의 관점에서 범용적으로 적용할 수 있는 국가 마스터데이터의 표준요소를 제안하고, 이를 통하여 효율적이고 표준적인 마스터데이

터의 필요성을 제안하는 것이다.

국내·외에서 이루어지는 마스터데이터 연구는 기업의 마스터데이터를 중심으로 이루어지고 있다. 마스터데이터에 관한 국외 선행연구는 마스터데이터를 구축한 기업을 중심으로 문제점을 지적하고 개선방안을 제시하는 연구가 주를 이룬다(Vilminko-Heikkinen and Pekkola 2017; Ofner et al. 2013; Haug and Arlbjorn 2011; Haapasalo 2011). 이 밖에 기존 연구의 한계를 염두에 두고 마스터데이터 구축의 필요성에 관해서도 연구가 이루어지고 있다. 이승민(2013)은 메타데이터 상호운용성에 관한 문제를 해결할 수 있는 대안으로 MDM 시스템 방식을 제안하였다. Wang, Ming, You(2009)는 효율적인 마스터데이터 관리 시스템 구축을 위해서 마스터데이터 구축이 선행되어야 한다는 필요성을 제기하고, 실제 마스터데이터 구축을 위한 프로세스를 제안하였다. 최영훈 외(2015)는 개인정보보호 종합포털에 등록된 자동차 관련 행정업무 및 생애주기별 복지 서비스 관련 데이터, 식품의약품 관련 데이터를 분석하여 기준데이터의 표준 코드(안)를 설계하였다. 이처럼 국내에서 이루어진 공공기관 마스터데이터에 대한 연구도 개인정보에 대한 연구로 제한점을 지니고 있다. 따라서 본 논문은 개인정보 이외에 공공기관에서 활용할 수 있는 마스터데이터를 연구하고자 한다.

1.2 연구의 범위와 방법

본 연구는 일차적으로 분석대상 마스터데이터의 범주를 기반시설로 한정하였다. 국내에서 이루어진 국가 마스터데이터 연구는 대부분 개

인정보 범주의 마스터데이터에 관련되어 있다. BRM 1레벨에 해당하는 15개 항목의 정책 분야 대부분에서 활용될 수 있는 데이터가 개인정보 마스터데이터이기 때문이다. 개인정보는 개인정보 보호법과 상충할 가능성이 항상 존재하므로, 개인정보를 실제 국가 마스터데이터로 구축하기 위해서는 법령 문제가 우선 해결되어야 한다. 반면, 기반시설 마스터데이터는 공공기관의 시설에 관한 부분으로 개인정보보다 적지만 대다수 기관에서 활용할 수 있는 데이터이다. 공공기관 정보는 기본적으로 국민에게 공개할 수 있는 데이터로 이용자 서비스 면에서 타 법령에 의한 제한점이 적은 데이터이다. 따라서 이 연구에서는 타 법령과 상충하지 않으면서 여러 기관에서 사용할 가능성이 높은 마스터데이터 범주가 기반시설이라고 판단하고, 기반시설에 대한 마스터데이터로 연구 대상을 한정하였다.

분석 대상은 공공데이터포털에서 제공하고 있는 표준 데이터셋으로 한정하였다. 일반적으로 마스터데이터를 구축하기 위해서 진행되는 연구는 개별적인 데이터를 분석하여 공통 표준요소를 추출하는 방식으로 진행되었다. 하지만 공공부문의 광범위한 데이터가 여러 웹사이트에 분산되어 제공된다는 점에서¹⁾ 개별 데이터를 종합적으로 분석하는 것이 어려웠다. 이에 대한 대안으로 본 연구에서는 법령에서 정의한 표준요소를 추출하여 정형화한 표준 데이터셋의 항목을 분석하고, 이를 기반으로 광의의 마스터데이터 표준요소를 구축한다.

본 연구는 문헌연구, 데이터 추출 및 분석, 개선방안 제안의 순서로 진행하였으며 연구의

세부적인 방법은 다음과 같다. 첫째, 문헌연구를 통해 데이터셋과 표준데이터의 법률적·행정적 정의를 조사하였으며, 마스터데이터와 관련한 국내·외의 선행연구를 분석하여 마스터데이터의 정의와 범주개념에 대하여 조사하였다. 둘째, 기반시설에 관한 표준 데이터셋 현황을 분석하고 개별적인 데이터셋 표준요소를 세 분야로 나누어 분석하였다. 셋째, 표준 데이터셋 표준요소를 비교분석하여 기반시설 데이터 간의 공통적인 항목을 추출하였다. 추출한 항목 이외에 기반시설 마스터데이터의 표준요소로 언급할 수 있는 요소를 제안하고, 그 결과를 종합하여 마스터데이터 표준요소를 제안하였다. 마지막으로 종합한 마스터데이터 표준요소를 국가에서 사용하고 있는 표준화 체계를 통하여 검증하였다.

2. 이론적 배경

2.1 데이터셋과 표준데이터의 정의

「공공기관의 정보공개에 관한 법률」(이하, 정보공개법)과 「공공데이터의 제공 및 이용 활성화에 관한 법률」(이하, 공공데이터법)에 따라 정부는 국민에게 공공데이터포털을 비롯하여 정보공개포털, 행정정보공동이용센터, 범정부 EA 포털 등의 다양한 플랫폼을 활용하여 공공데이터를 제공한다. 이 가운데 공공데이터포털은 데이터셋을 기준으로 관리되고 있다. 데이터셋은 「공공데이터 개방 표준」에서는 '주차장,

1) 공공데이터포털에서만 2만 건 이상의 데이터셋을 제공하고 있고, 정보공개 홈페이지에는 11만 7천 건 이상의 즉시 원문을 열람할 수 있는 원문정보를 공개하고 있다.

도시공원, 약국, 병의원 등 공공데이터 분야별 데이터의 집합으로 정의하고, 『공공데이터 관리 지침』에서는 기록을 관리하는 원천데이터 내 유사데이터의 묶음 및 집합으로 정의하고 있다. 한국기록학회(2008)에서는 데이터세트를 컴퓨터가 처리하거나 분석할 수 있는 형태로 존재하는 관련 정보의 집합체로 정의한다. 이 연구에서는 이를 기반으로 데이터세트를 기록정보활용을 위하여 데이터를 제공하는 다양한 형식 가운데 하나로서 법령 등에 따라 표준화될 수 있는 데이터의 집합으로 정의한다.

데이터세트의 표준화에 관한 법령은 『공공데이터법』 제23조와 『공공데이터 개방 표준』, 『공공데이터 관리 지침』이 있고, 관련 법령에 따라 정립된 공공데이터 개방표준은 2017년 10월을 기준으로 93개 항목이다. 공공기관은 개방표준으로 정해진 93개 분야의 개별 데이터세트를 작성할 경우, 해당 분야의 표준에서 기술하고 있는 제공항목 기준을 바탕으로 데이터세트를 작성해야 한다. 2017년 10월 『공공데이터 개방 표준』의 개정 전까지 개방표준에 기반을 두어 표준데이터 항목을 구성하는 주체는 공공데이터 활용지원센터로, 개별적인 데이터세트건을 종합하여 표준데이터를 구축하였다. 하지만 표준의 개정 후 구축된 시스템을 통해 각 기관이 직접 표준데이터 항목을 구성할 수 있게 되었다. 공공데이터 개방표준에 따라 각 기관에서 작성된 데이터파일은 기관이 직접 개방표준데이터점검서비스를 통해 개방표준에 일치하는 지 점검할 수 있다. 점검결과가 적합한 데이터파일은 공공데이터포털(관리자접속)을 통해 이미 등록된 표준데이터에 추가 등록하거나 새로운 표준데이터를 생성한다. 2017년 10월을 기준으로 구축

된 표준데이터는 58건으로, 나머지 35건의 표준데이터로 구축되지 못한 개방표준 항목은 개별적인 데이터세트로 제공되고 있다.

데이터세트는 개별 기관에서 작성한 개별 데이터 한 건을 의미하는 것으로 공공데이터포털에서 오픈 API와 파일 데이터 항목으로 제공된다. 공공데이터 개방표준에서 정의하고 있는 93건의 유형에 한해 여러 기관의 데이터세트를 통합한 파일이 표준데이터이다. 개방표준에서 정의하고 있는 93개 데이터세트 범주 이외의 데이터세트는 표준데이터로 구축될 수 없다.

2.2 마스터데이터의 정의와 특징

기존의 연구는 마스터데이터에 대한 분석적 관점으로 수행한 연구가 상대적으로 작고 마스터데이터 관리 시스템 관점에서 마스터데이터를 부분적으로 분석한 연구가 상대적으로 많았다. Oracle(2011)의 정의에 따르면 MDM 시스템을 활용하고 있는 조직의 데이터는 크게 트랜잭셔널데이터(Transaction Data)와 마스터데이터, 분석데이터(Analytical Data)로 구분할 수 있다. 이 중 마스터데이터는 국내에서 기준데이터라는 명칭으로 사용하며, 그 세부적인 정의와 범주는 구체적으로 정리된 사항이 없고 연구자별로 다르게 정의하고 있다. 2006년부터 2016년까지 이루어진 마스터데이터의 정의는 기본적인 회사 기준데이터로 귀결되었으며 그 기준데이터는 기록학적 의미와 함께 활용적 의미로 사용하기 위한 데이터로 정의하고 있다. 대부분의 정의는 기업의 관점에서 이루어졌으나 예외적으로 최영훈 외(2015)는 정부 관점에서 국가운영체계에서 사용하는 마스터데이터

에 대해 그 중요성을 설명하고 있었다. 국내외 학자별 마스터데이터의 정의 및 범주를 종합하면 <표 1>과 같다.

<표 1>의 마스터데이터 정의에 따르면 마스터데이터는 여러 업무에서 공통으로 참조될 수 있는 중요 데이터로 쉽게 변화하지 않는 특성을 갖는다. 이 정의를 공공부문에 적용하면, 마스터데이터는 국가의 여러 부처에서 참조하여 사용할 수 있는 공통적인 데이터이다. 기관 전체에서 사용하는 마스터데이터의 특징을 고려하여 마스터데이터의 정의와 범주를 세부적이고 명확하게 정의할 경우 마스터데이터의 효용성 또한 제한적으로 정의된다. 즉, 마스터데이터는 기관의 특성을 고려하는 기준이므로 모든 기관에서 사용되는 마스터데이터의 정의와 범주를 세부적으로 한정하기에 한계가 있다. 특

히 국가 마스터데이터의 범주는 기업에서 사용되는 마스터데이터 범주와 차이가 있으므로 구축 단계에서 업무 분석을 통하여 국가 운영 실정에 맞는 범주를 선정하여야 한다.

마스터데이터를 기존의 다른 데이터 통합 방식(예를 들어, 명감, 디렉토리, 메타데이터)과 비교하면 개념을 명확히 할 수 있다. 명감은 기관별로 임의 생산한 자료로, 국가 통제를 필요로 하지 않은 자료이다. 반면, 마스터데이터는 국가차원의 전자자원의 공통요소를 도출하여 활용하는 중심자원으로, 선택사항인 명감과 달리 필수사항으로서의 성격을 지니고 있다. 이러한 관점에서 마스터데이터는 전자형태로 이루어진 정부 생산·관리 데이터를 공유하여 활용할 수 있는 기준데이터이다. 기반시설 데이터를 활용하는 기관의 경우 마스터데이터의 내

<표 1> 마스터데이터의 정의 및 범주

	연구자	년도	정의 및 범주
정의	박진경, 김미수	2016	기업이 행하는 다양한 업무에서 기본적으로 다루는 공통된 데이터를 하나의 중앙 데이터베이스로 모아 둔 것
	Rivas 외 4인	2016	하나의 기관이 그들의 업무 활동을 개발할 때 업무 도메인의 기본적인 지식을 결정하는 개념
	이승민	2013	기업 내에서 체계적인 데이터 관리를 위한 변화의 결과로 나타나게 된 것. 모든 기업 활동에 있어서 기본이 되는 자료이며 관련된 정보를 관리하는데 사용될 수 있는 통제 기능을 수행
	Ofner 외 3인	2013	회사 업무 활동의 기본이 되는 중요한 업무 요소를 명시화하는 기준
	Berson과 Dubov	2007	주요 업무 활동을 위하여 전사적 "기록시스템(system of record)"으로 정화, 합리화, 통합된 데이터
	조현구	2006	업무 처리에 자주 같은 형태로 사용되는 정보로 여러 개의 제어데이터(control data)와 관련 정보를 가지면서 장기간 동일하게 유지되는 데이터
범주	최영훈, 김태진, 서승현, 정세희	2015	고객, 제품, 공급자, 구매자, 직원, 재무, 정책
	Cervo, Allen	2011	고객, 제품, 공급자, 파트너, 직원, 재료
	Oracle	2011	고객, 공급업체, 장소, 계좌, 자산, 제품
	Loshin	2009	고객, 직원, 벤더, 공급업체, 파트, 제품, 위치, 연락 메커니즘, 프로파일, 계좌, 계약, 정책
	Wang, Ming, You	2009	고객, 공급업체, 장소, 제품, 물품목록

용을 활용하여 각 기관별로 데이터를 작성할 수 있다. 이러한 방식의 장점은 주요 데이터를 마스터데이터로 공유함으로써 각 기관별로 생성되는 메타데이터 간의 차이점을 방지할 수 있으며, 데이터를 수집하는 데 드는 시간과 비용을 일정부분 경감시킬 수 있다는 점이다. 또한 각 기관에서 제공되는 정보의 차이를 줄여 서비스 시 이용자의 혼란을 방지할 수 있다는 점도 주요 장점이다.

데이터의 품질적인 면과 마스터데이터를 정의한 국내의 연구들을 정리하면 마스터데이터의 특징은 크게 6가지로 구분할 수 있다.

첫째, 마스터데이터는 기관별로 세부적인 데이터와 연계될 수 있지만, 마스터데이터 자체의 기본속성만큼은 장기간 동일하게 유지되어야 한다(Dreibelbis et al. 2008). 마스터데이터의 속성이 빈번하게 변경될 경우 전체적인 관리비용이 증가하게 되고, 마스터데이터 구축의 필요성이 저하될 수밖에 없다.

둘째, 마스터데이터는 업무상 빈번하게 사용되는 중요한 데이터여야 한다(Loshin 2009). 마스터데이터는 모든 활동에 있어서 기본이 되는 자료로, 회사에서 가장 중요하고 필수적인 데이터이다(Rivas et al. 2016). 이러한 자료는 업무상 빈번하게 사용되며 마스터데이터로 구축될 경우 그 활용성과 경제성이 높은 데이터이다.

셋째, 범정부 차원에서 전반적으로 사용되어야 한다. 기업에서는 기업 전반에 걸쳐 사용되어야 하는 것처럼(Loshin 2009), 정부에서는 범정부 차원에서 다양한 분야에 걸쳐 여러 방면으로 사용되어야 한다. 여러 부문에서 함께 사용되는 데이터가 마스터데이터로 구축될 경우, 여러 기관에서 중복하여 생성 및 관리하던 데이터의 관

리비용을 절감할 수 있고 일관성 있는 데이터 관리가 가능하게 될 것이다.

넷째, 마스터데이터는 그 자체로 완전해야 한다(Ofner et al. 2013; Dreibelbis et al. 2008). 기록관리의 측면에서 완전성은 기록이 생산되었을 때 필요한 정보가 모두 포함되어 있음을 보증하는 속성을 의미한다(한국기록학회 2008). 기록이 완전성을 가지기 위해서는 그 기록만으로 해당하는 내용을 완전히 이해할 수 있어야 하며, 이를 위해 관련된 모든 요소가 포함되어야 한다.

다섯째, 마스터데이터는 정확성을 가지고 있어야 하며(Ofner et al. 2013), 마스터데이터의 정확성을 확보하기 위해서는 식별자를 통한 항목별 구분이 가능하여야 한다(Dreibelbis et al. 2008). 마스터데이터가 여러 기관에서 중복되어 관리될 경우 같은 데이터가 서로 다른 값을 가지고 관리될 수 있다. 이 경우 해당 데이터의 신뢰성이 떨어지며, 데이터를 확인하는데 드는 비용과 서비스의 단절성 면에서 문제가 발생한다. 이 경우 식별자를 통하여 마스터데이터 내에 중복된 자료를 쉽게 발견할 수 있다.

여섯째, 마스터데이터는 일관성을 유지하기 위하여 정형화된 구조를 사용해야 한다(Dreibelbis et al. 2008). 마스터데이터는 여러 기관에서 사용할 수 있는 데이터를 모아둔 데이터의 집합으로, 하나의 범주에 많은 양의 데이터를 보관한다. 따라서 데이터 자체가 간결하여야 하며, 기관별로 내용을 추가할 때 다른 데이터와 동일한 방식으로 작성되어야 한다.

2.3 기반시설의 정의와 범주

본 논문은 마스터데이터 구축의 범주를 기반

시설로 한정하였다. 기반시설에 대한 데이터셋을 추출하기 위해서는 기반시설에 포함되는 세부 시설 범주가 명시되어야 한다. 이를 위하여 「국토의 계획 및 이용에 관한 법률」(이하, 국토계획법) 제2조 6항, 「국토계획법」 시행령과 국토계획법의 하위법령인 국토교통부령 「도시·군계획시설의 결정·구조 및 설치기준에 관한 규칙」법령에 기반을 두어 세부적인 기반시설의 범주를 <표 2>로 구체화하였다. 표의 시설 명칭은 법령 명칭과 동일하게 표기하였다.

3. 기반시설의 표준 데이터세트 요소 분석

3.1 기반시설 표준 데이터세트 현황

표준 데이터세트 현황은 행정업무용 표준인 「공공데이터 개방 표준」을 참고하여 분석하였다. 93건의 데이터세트별 개방표준 중 표준데이터로 구축되지 못한 38건의 표준 데이터세트 요소도 명확하게 규정된 속성정보를 갖는다. 속성

<표 2> 기반시설의 세부적인 시설 구분

법률	대통령령	법률	대통령령	
교통시설	도로	공공·문화 체육시설	공공청사	
	철도		문화시설	
	항만		체육시설	
	공항		연구시설	
	주차장		사회복지시설	
	자동차정류장		공공직업훈련시설	
	궤도		청소년수련시설	
	운하		방재시설	하천
	자동차 및 건설기계검사시설			유수지
	자동차 및 건설기계운전학원			저수지
공간시설	광장	방화설비		
	공원	방풍설비		
	녹지	방수설비		
	유원지	사방설비		
	공공공지	방조설비		
유통·공급시설	유통업무설비	보건위생시설	화장시설	
	수도공급설비		공동묘지	
	전기공급설비		납골시설·자연장지	
	가스공급설비		장례식장	
	열공급설비		도축장	
	방송·통신시설		종합의료시설	
	공동구		환경기초시설	하수도
시장	폐기물처리시설			
유류저장 및 송유설비	수질오염방지시설			
공공·문화 체육시설	학교	폐차장		
	운동장			

정보가 정의된 개방표준은 자료의 충분한 수집이 이루어질 경우, 자료를 보유하고 있는 개별 기관에서 직접 개방표준데이터점검서비스를 통해 표준데이터를 구축할 수 있다. 따라서 본 논문에서는 표준데이터로 구축되지 못한 개방표준데이터라도 근 시일 내 표준데이터로 구축될 가능성이 높다고 판단하여 모든 개방표준을 대상으로 분석하였다.

현재 국가에서 제공되고 있는 93건의 데이터 세트 중 법률에 기반을 두고 규정한 <표 2> 기반시설의 세부적인 시설 구분에 해당하는 정보를 다루는 데이터세트를 추출하였다. 이 중 마스터데이터로서 지녀야 하는 기본적인 특성에 해당하지 않는 쉽게 변화하는 데이터세트를 제외하였다. 결과적으로 본 논문에서는 마스터데

이터의 특징을 지닌 기반시설 관련 13개 표준데이터를 대상으로 분석을 실시하였다. 각 표준별로 제공하고 있는 파일과 분류체계를 정리한 결과는 <표 3>과 같다.

<표 3>의 표준 데이터세트 명칭은 『공공데이터 개방 표준』에서 정의하고 있는 표준 명칭을 그대로 차용한 것이다. 이는 실제로 공공분야에서 사용하고 있는 데이터와 체계를 현상 그대로 사용함으로써 제시하는 안에 대한 실제 적용 가능성을 확보하기 위함이다. 개방대상기관 항목과 표준데이터 분류체계 항목은 공공데이터 개방 표준을 정리한 것이고, 표준데이터 제공여부와 개별 파일데이터 분류체계는 공공데이터포털의 데이터를 재구성한 것이다.

13개 표준 중 표준데이터로 제공되는 항목은

<표 3> 기반시설 표준 데이터세트의 표준데이터 제공여부 및 분류체계

(2017년 10월 기준)

기반시설 범주	표준 데이터세트 명칭	개방대상기관	표준데이터 분류체계	파일데이터 개수	개별 파일데이터 분류체계
공원	도시공원정보	국토교통부, 지방자치단체	문화관광	330	국토관리, 공공행정, 사회복지, 문화관광, 교통물류, 환경기상, 농축수산
유원지	휴양림	산림청, 지방자치단체	문화관광	138	공공행정, 문화관광, 농축수산
유원지	야영(캠핑)장	지방자치단체	문화관광	121	문화관광, 농축수산
시장	전통시장	중소기업청, 지방자치단체	문화관광	326	국토관리, 공공행정, 재정금융, 산업고용, 문화관광, 교통물류
철도	도시철도 역사정보	도시철도공사, 지방자치단체 등	교통물류	11	교통물류
자동차 및 건설기계 검사시설	자동차검사소	국토교통부, 교통안전공단, 지방자치단체	교통물류	37	공공행정, 교통물류
폐기물처리시설	재활용센터	지방자치단체	사회복지	83	공공행정, 사회복지, 환경기상
체육시설, 문화시설	공공시설개방	지방자치단체	공공행정	311	교육, 국토관리, 공공행정, 사회복지, 문화관광
문화시설	도서관	문화체육관광부, 지방자치단체	교육	361	교육, 국토관리, 공공행정, 문화관광
문화시설	박물관·미술관	문화체육관광부, 지방자치단체	문화관광	185	문화관광
학교	초·중등학교위치	교육부, 한국교육개발원	교육	1	교육
사회복지 시설	건강증진센터	지방자치단체	보건의료	131	보건의료
사회복지 시설	치매센터	보건복지부, 지방자치단체	보건의료	104	보건의료

2017년 5월을 기준으로 9개였으나, 2017년 10월을 기준으로 13개를 전부 표준데이터로 제공되고 있다. 구축 주체가 공공데이터활용지원센터라는 단일 기관이었을 때는 이미 제공되고 있는 데이터가 일정 수준이 넘을 경우에만 공공데이터포털에서 데이터를 종합하여 구축하였다. 즉, 데이터세트가 일정 수준에 이르기 전까지는 하나의 표준데이터로 종합하여 관리되지 못하였다. 하지만 표준데이터 구축 주체가 단일 기관에서 데이터를 실제로 구축하는 복수의 개별 기관으로 변화함에 따라 표준데이터 구축이 상대적으로 용이하게 됨으로써 향후 표준데이터 수는 지속적으로 증가할 것이다.

표준데이터 구축이 과거에 비해 수월하게 변화되었음에도 불구하고 개별 데이터세트는 표준데이터로 통합되기 이전까지 개별 기관에서 작성·관리하는 제약점을 갖고 있다. 따라서 표준에서 명시하고 있지 않은 부분에 대해서는 개별 기관의 주관적인 기준으로 관리되는 특징을 갖는다. 대표적으로 표준데이터의 분류체계는 소속기관에서 개별적으로 구축하는 파일데이터의 특성에 따라 다른 분류체계에 속할 수 있다. 하나의 개방표준이 존재함에도 구축되는 주체가 다르고 처음부터 종합·관리되지 않는다는 특성 때문에 제공하는 데이터의 일관성에 제한이 생길 수 있다. 실제 <표 3>과 같이 표준데이터를 구축하는 개별 파일데이터가 많을수록 분류체계가 다양하게 나타난다. 교육부가 개방을 담당하는 초중등학교 위치 표준데이터를 제외한 나머지 12개 표준데이터는 모두 지방자치단체를 개방대상기관으로 지정하고 있다. 교육부는 전체 지역을 대상으로 데이터를 한 곳에서 관리하고 있기 때문에 파일데이터

개수도 적고 분류체계도 동일하게 나타난다. 하지만 대다수의 표준데이터 개방대상기관은 지방자치단체에서 개방하고 있어 실제적으로 관련 데이터는 지방자치단체 등에서 개별적으로 생성·관리되고 있다.

표준 데이터세트는 하나의 참고표준으로써 제공될 수 있지만, 전체 표준으로서 작용하기 위해서는 제한점이 존재한다. 분산 관리되고 있는 데이터 간의 일치를 위해서는 하나의 관리체계를 지정하고 해당 체계 내에서 데이터를 통합하여 관리하여야 한다. 따라서 현재 작성되어 사용하고 있는 데이터를 무시하고 새로운 체계를 구축하기 위해서는 상대적으로 많은 노력과 비용이 필요하다. 즉 여러 기관에서 사용되는 데이터 체계의 정합성을 유지하고 변경하는 시간이 상대적으로 많이 소요될 수 있다. 따라서 본 논문에서는 기관 내에서 사용하고 있는 데이터 체계인 개방표준을 고려하여 데이터의 마스터데이터화를 진행한다.

3.2 기반시설 표준 데이터세트 분석

기반시설은 각기 고유한 목적과 특성을 보인 다양한 유형의 시설을 포함하고 있다. 같은 기반시설 범주에 속하는 표준데이터라도 특정 시설에서만 사용할 수 있는 서로 다른 표준항목을 갖는다. 그럼에도 불구하고 기반시설 범주에 속하는 시설들이 공공기관으로서의 공통된 특징을 지니고 있으며, 고유 특징 이외의 일반 정보가 다른 기관에서도 사용할 수 있기 때문에 기반시설의 공통 항목이 의미를 지닐 수 있다. 이러한 특성을 고려하여 각 표준데이터에서 제시하고 있는 속성을 해당 시설에서만 사용가능한

항목, 비슷한 시설에서 사용 가능한 항목, 모든 시설에서 사용 가능한 항목으로 구분하였다. 이 가운데 해당 시설에서만 사용 가능한 항목은 다른 기반시설에서는 사용할 수 없는 해당 시설만의 특성을 보인다. 이 항목은 전반적인 기반시설에 관련된 마스터데이터를 구축할 때 사용하기 어려운 항목이기 때문에 본 논문에서는 분석 대상에서 제외하였다.

유사 시설에서 사용 가능한 항목은 특정 범주에 해당하는 시설에서 사용할 수 있는 표준항목을 정리하고 있다. 예를 들어 특정 기관의 세부 시설 정보를 비롯하여 수용인원 관련 정보, 주차장 관련 정보, 홈페이지 정보, 전문 인력 정보,

운영일 정보, 관람료 정보 등이 이에 해당한다. 이러한 표준요소는 특정 시설에서만 사용 가능한 항목이 아닌 여러 기관에서 사용할 수 있는 항목이라는 점에서 분석 대상으로 선정하였다. 13개 표준 데이터세트의 표준요소 중 비슷한 시설에서 사용 가능한 항목을 정리하면 <표 4>와 같다.

모든 시설에서 사용 가능한 항목은 특정 시설 이외에 다른 시설에서도 활용 가능한 항목으로 마스터데이터로 구축될 가능성이 가장 높은 항목이다. 예를 들어 위도와 경도의 위치정보는 시설이 가져야 하는 가장 기본적이고 필수적인 요건이다. 또한 모든 시설에서 사용 가

<표 4> 비슷한 유형의 기반시설에서 사용 가능한 표준요소

표준 데이터세트 명칭	비슷한 유형의 시설에서 사용가능한 항목
도시공원정보	공원면적, 공원보유시설(운동시설), 공원보유시설(유희시설), 공원보유시설(편익시설), 공원보유시설(교양시설), 공원보유시설(기타시설), 지정고시일
휴양림	시도명, 휴양림면적, 수용인원수, 숙박가능여부, 주요시설명, 홈페이지주소
야영(캠핑)장	부지면적, 건축연면적, 1일최대수용인원수, 주차장면수, 편의시설, 안전시설, 기타부대시설, 이용시간, 이용요금
전통시장	홈페이지주소, 공중화장실보유여부, 주차장보유여부, 개설년도
도시철도 역사정보	한자역사명, 영문역사명
자동차검사소	자동차검사소유형, 운영시간, 부지면적, 건물면적, 검사기술인력수
재활용센터	재활용센터운영구분, 면적, 설립년월, 평일운영시작시각, 평일운영종료시각, 공휴일운영시작시각, 공휴일운영종료시각, 휴무일정보, 홈페이지주소
공공시설개방	개방장소명, 개방시설유형구분, 휴관일, 평일운영시작시각, 평일운영종료시각, 주말운영시작시각, 주말운영종료시각, 유료사용여부, 수용가능인원수, 면적, 부대시설정보, 시설사진정보, 담당부서명, 홈페이지주소
도서관	시도명, 시군구명, 휴관일, 휴관일, 평일운영시작시각, 평일운영종료시각, 토요일운영시작시각, 토요일운영종료시각, 공휴일운영시작시각, 공휴일운영종료시각, 부지면적, 건물면적, 홈페이지주소
박물관·미술관	박물관·미술관구분, 운영홈페이지, 편의시설정보, 평일관람시작시각, 평일관람종료시각, 공휴일관람시작시각, 공휴일관람종료시각, 휴관정보, 어른관람료, 청소년관람료, 어린이관람료, 관람료기타정보
초·중등학교위치	설립일자
건강증진센터	운영시작시각, 운영종료시각, 휴무일정보, 건물면적, 의사수, 간호사수, 사회복지사수, 영양사수, 기타인력현황
치매센터	설립년월, 건축물면적, 부대시설정보, 의사인원수, 간호사인원수, 사회복지사인원수, 기타인원현황, 운영위탁일자

〈표 5〉 모든 기반시설에서 사용 가능한 표준요소

표준 데이터세트 명칭	모든 시설에서 사용가능한 항목
도시공원정보	공원명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 관리기관명, 전화번호, 데이터기준일자
휴양림	휴양림명, 소재지도로명주소, 관리기관명, 휴양림전화번호, 위도, 경도, 데이터기준일자
야영(캠핑)장	야영(캠핑)장구분, 위도, 경도, 소재지도로명주소, 소재지지번주소, 야영장전화번호, 관리기관전화번호, 관리기관명, 데이터기준일자
전통시장	시장명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 전화번호, 데이터기준일자
도시철도 역사정보	역사명, 역위도, 역경도, 운영기관명, 역사도로명주소, 역사전화번호, 데이터기준일자
자동차검사소	자동차검사소명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 검사소전화번호, 관리기관전화번호, 관리기관명, 데이터기준일자
재활용센터	재활용센터명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 운영기관명, 운영기관전화번호, 운영기관대표자명, 관리기관전화번호, 관리기관명, 데이터기준일자
공공시설개방	개방시설명, 소재지도로명주소, 소재지지번주소, 관리기관명, 사용안내전화번호, 위도, 경도, 데이터기준일자
도서관	도서관명, 소재지도로명주소, 운영기관명, 도서관전화번호, 위도, 경도, 데이터기준일자
박물관·미술관	시설명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 운영기관전화번호, 운영기관명
초·중등학교위치	학교명, 소재지지번주소, 소재지도로명주소, 교육지원청명, 위도, 경도, 데이터기준일자
건강증진센터	건강증진센터명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 운영기관전화번호, 운영기관명, 관리기관전화번호, 관리기관명, 데이터기준일자
치매센터	치매센터명, 소재지도로명주소, 소재지지번주소, 위도, 경도, 운영기관명, 운영기관대표자명, 운영기관전화번호, 관리기관전화번호, 관리기관명, 데이터기준일자

능한 항목도 위도와 경도처럼 항구적인 것 이외에 정책의 변화에 따라 데이터가 달라지는 항목도 있다. 관리기관명은 정부조직법과 같이 상급단체의 명칭이 법적으로 변화할 때 항구성을 유지하지 못한다. 단, 이와 같은 항목들은 기록전거데이터로써 처리할 수 있으면 항구성을 유지하는 공통항목으로 처리할 수 있다. 13개 표준 데이터세트의 표준요소 중 모든 시설에서 사용 가능한 항목만을 정리하면 〈표 5〉와 같다.

마스터데이터를 구축하기 위해서는 이미 존재하는 데이터의 체계를 고려하여 여러 기관에서 사용할 수 있는 체계를 구축하여야 한다. 기반시설에 관한 표준 데이터세트의 세부적인 표준요소를 세 항목으로 구분한 결과, 비슷한 시설에서 사용 가능한 항목과 모든 시설에서 사용 가능한 항목에 속한 표준요소가 마스터데이

터로서 활용될 수 있기 때문에 세부적인 분석을 진행하였다.

4. 기반시설 마스터데이터 표준요소 구축

4.1 기반시설 표준 데이터세트 비교

13건의 기반시설 관련 표준 데이터세트의 제 공항목 중 모든 시설에서 사용 가능한 항목명을 살펴보았다. 항목명이 정확히 일치하지 않을 수도 있지만, 대다수 표준 데이터세트가 명칭, 소재지 도로명주소, 소재지 지번주소, 위도, 경도, 데이터기준일자, 관리기관 및 운영기관 정보를 포함하고 있다. 이를 세부적인 표준별

로 정리하면 <표 6>과 같다. 이 표는 표준에서 제시하고 있는 표준명과 항목명을 그대로 차용하였다. 필수선택은 해당 제공항목을 필수로 기입해야 하는지 아닌지를 나타내는 것으로, 공공데이터 개방 표준에 명시하고 있는 필수선택 사항을 도형으로 변환하여 표기한 것이다.

<표 6>에 따르면 분석대상인 13건의 개방표준이 모두 가지고 있는 항목명은 명칭과 소재지 도로명주소, 위도, 경도, 그리고 데이터기준 일자의 총 5개 항목이다. 위도와 경도, 관리기관명칭은 특정 분야에서 선택요소이지만 8개 이상의 분야에서 필수요소로 조사되었다. 이중 전화번호와 명칭은 두 가지 이상의 항목을 구분하여 기입하고 있다.

첫째, 전화번호는 13개 데이터세트에서 운영기관 전화번호와 관리기관 전화번호로 나뉘어 기입하고 있다. 두 전화번호 중 하나도 사용하

지 않고 있는 기관은 전통시장과 초중등학교위치 데이터세트이며, 나머지 데이터세트는 두 전화번호를 모두 제공하거나 둘 중 하나만 제공하고 있다. 전통시장의 경우 전화번호 항목을 제시하고 있지만, 이 전화번호는 운영기관 전화번호나 관리기관 전화번호로 명확하게 구분하기 힘든 항목이다. 전화번호 항목을 사용하고 있는 11개 데이터세트 중 6개 데이터세트에서는 운영기관 전화번호와 관리기관 전화번호를 모두 기입하고 있다. 4개 데이터세트에서는 운영기관 전화번호 항목만을 사용하고 있다. 또한, 운영기관 전화번호를 사용하지 않고 관리기관 전화번호만 사용하고 있는 표준 데이터세트는 도시공원 데이터세트뿐이다. 도시공원의 전화번호는 공원의 관리부서 전화번호로 관리기관 내의 특정 부서 전화번호이다. 이처럼 운영기관 전화번호와 관리기관 전화번호를 모두 가지고 있는

<표 6> 기반시설 관련 표준 데이터세트의 공통 제공항목

	명칭	도로명주소 소재지	지번주소 소재지	위도	경도	전화번호 운영기관	명칭 운영기관	전화번호 관리기관	명칭 관리기관	기준일자 데이터
도시공원정보	●	●	▲	●	●			▲	▲	●
휴양림	●	●	●	●	●	▲			●	●
야영(캠핑)장	●	●	▲	●	●	●		●	●	●
전통시장	●	●	▲	▲	▲					●
도시철도 역사정보	●	●		●	●	●			●	●
자동차검사소	●	●	▲	●	●	●		▲	●	●
재활용센터	●	●	▲	●	●	▲	●	●	●	●
공공시설 개방정보	●	●	▲	●	●	▲			●	●
도서관	●	●	●	▲	▲	▲			▲	●
박물관·미술관정보	●	●	▲	●	●	●	●	●	●	●
초·중등학교 위치	●	●	●	●	●				●	●
건강증진센터	●	●	▲	●	●	●	▲	▲	●	●
치매센터	●	●	▲	●	●	●	●	▲	●	●

●의 표시는 주관기관이나 법적인 배경의 여부와 관계없이 모든 분야에서 동일하게 존재하는 필수항목을 의미한다.

▲의 표시는 기관의 특성과 법적인 배경에 따라 선택적으로 나타나는 선택항목을 의미한다.

데이터세트도 있으므로 단순히 전화번호 항목으로 통합하기에 어려움을 갖는다. 왜냐하면, 두 기관의 전화번호 중 하나를 선택하여야 하는 문제가 생길 수 있고, 추후 다른 기관에서 이 데이터를 이용할 때에 전화번호 항목의 명확한 의미를 파악하기 어렵기 때문이다.

둘째, 운영기관 혹은 관리기관 명칭 항목 중 하나도 사용하고 있지 않은 데이터세트는 전통시장 데이터세트뿐이다. 전통시장의 경우 관리가 이루어지는 시설이 아니므로 관리기관이나 운영기관 명칭을 가지고 있지 않다. 전화번호 항목을 제시하고 있지 않은 초중등학교 위치 데이터세트의 경우에도 설립 형태에 무관하게 관리기관에 해당하는 교육청과 교육지원청 명칭을 공개하고 있다. 운영기관명칭과 관리기관명칭을 모두 사용하고 있는 데이터세트는 재활용센터, 박물관·미술관정보, 건강증진센터 및 치매센터이다. 4개 기반시설의 경우 운영기관과 관리기관을 명확히 구분할 수 있으므로 관리기관과 운영기관의 명칭을 모두 사용하고 있다.

모든 기반시설을 대상으로 공통요소를 추출할 경우 일관된 특성을 보인 기관들의 요소를 추출하는 한계가 있다. 따라서 일부 비슷한 특성을 공유하고 있는 기관이 공통으로 가지고 있는 요소에 기반을 두어 기반시설을 몇 가지 범주로 분류하여 분석할 필요가 있다. 범주로 사용할 수 있는 특성은 3장 2절에서 분석한 세 항목의 표준 요소 중 비슷한 시설에서 사용 가능한 항목으로 기준을 설정할 수 있다. 예를 들어 숙박 가능 여부와 전문인력 존재 여부, 유료 사용 여부, 운영일의 존재 여부, 세부 보유시설 기재 여부 등을 범주로 활용할 수 있다.

세부적으로 숙박 가능 여부와 전문인력 존재

여부의 경우 13개 기반시설을 대상으로 하였을 때 일부의 시설에서만 활용할 수 있다. 숙박 가능 여부 항목의 경우 휴양림과 야영(캠핑)장에서만 사용할 수 있고, 전문인력 존재 여부 항목은 전문 인력이 있어야 하는 자동차검사소를 비롯한 건강증진센터, 치매센터에서만 활용할 수 있다. 반면 운영일의 존재 여부의 경우 13개의 기반시설 관련 데이터세트 중 가장 많은 7개 시설에서 사용하고 있는 항목명이다. 따라서 본 논문에서는 공통항목을 가장 많이 추출할 수 있는 운영일을 기준으로 범주를 재설정하였다. 운영일에 관련된 휴무일, 휴관일, 운영 시작 및 종료시각 항목 중 하나라도 지정하고 있는 표준 데이터세트를 선정하였고, 해당 표준데이터의 제공 항목은 <표 7>과 같다.

분석 대상이 되는 7개 데이터세트에서 모두 확인할 수 있는 항목은 명칭을 비롯하여 소재지 도로명주소, 소재지지번주소, 위도, 경도, 운영기관전화번호, 관리기관명칭, 운영시간, 데이터기준일자이다. 이 가운데 공통 요소로 나타나지 않았던 항목은 소재지지번주소를 비롯하여 운영기관전화번호, 관리기관명칭, 운영시간 항목이다. 이 중 소재지지번주소는 소재지도로명주소로 대체할 수 있고 도서관 데이터세트를 제외하고 나머지 데이터세트에서 선택항목으로 제공되고 있으므로 공통항목에서 제외하였다.

운영기관전화번호와 관리기관명칭은 모든 데이터세트에서 확인할 수 있고 다른 항목과 중복되지 않기 때문에 공통항목으로 제시할 수 있다. 휴무일의 경우 7개 데이터세트 중 야영(캠핑)장과 자동차검사소만 휴무일정보 항목 제공하지 않는다. 하지만 두 데이터세트 모두 운영시간을 통해 휴무일을 확인할 수 있고, 자동

〈표 7〉 운영일이 있는 기반시설 표준 데이터세트의 제공항목

	야영 (캠핑)장	자동차 검사소	재활용 센터	공공시설 개방정보	도서관	박물관 미술관	건강증진 센터
명칭	●	●	●	●	●	●	●
소재지도로명주소	●	●	●	●	●	●	●
소재지지번주소	▲	▲	▲	▲	●	▲	▲
위도	●	●	●	●	▲	●	●
경도	●	●	●	●	▲	●	●
운영기관전화번호	●	●	▲	▲	▲	●	●
운영기관명칭			●			●	▲
관리기관전화번호	●	▲	●			●	▲
관리기관명칭	●	●	●	●	▲	●	●
휴무일정보/휴관일			▲	●	●	●	●
운영시간/이용시간	▲	●	●	●	●	●	●
유료사용여부	▲			●		●	
홈페이지주소			▲	▲	▲	▲	
데이터기준일자	●	●	●	●	●	●	●

●의 표시는 주관기관이나 법적인 배경의 여부와 관계없이 모든 분야에서 동일하게 존재하는 필수항목을 의미한다.

▲의 표시는 기관의 특성과 법적인 배경에 따라 선택적으로 나타나는 선택항목을 의미한다.

차검사소의 경우 평일과 토요일, 공휴일의 운영 시간을 덧셈(+) 표시로 구분하여 하나의 항목으로 제공하고 있다. 가장 세분화 수준이 낮은 야영(캠핑)장 데이터세트에서는 단순히 이용 시간만을 기입하고 있다. 나머지 5개 표준 데이터세트는 휴무일정보와 운영시간 정보를 모두 제공하고 있지만, 세분화 수준에서 조금씩 차이를 보인다.

위의 분석 결과를 통하여, 전체 기반시설을 대상으로 마스터데이터를 구축하기에는 한계가 존재한다는 것을 파악할 수 있다. 이는 다양한 시설을 대상으로 할 경우 시설이 가지고 있는 특징을 충분히 표기하지 못하기 때문이다. 하지만 일부 공통적인 특징을 공유하고 있는 기반시설을 범주를 지정하여 분석할 경우, 기반시설은 공통적인 메타데이터를 가질 수 있었다. 하지만 위에서 추출한 모든 공통항목이 마

스터데이터로 활용될 수 있는 것은 아니다. 쉽게 변화하는 속성이거나 범정부적으로 사용될 수 없는 마스터데이터로서의 특징에 반하는 항목의 경우 공통적인 메타데이터이지만 마스터데이터의 항목이라고 할 수 없다. 따라서 분석한 결과에서 마스터데이터의 특성에 부합하는 항목만을 마스터데이터 항목으로 제한할 필요가 있다.

4.2 마스터데이터 표준요소 설계

4.2.1 기반시설 마스터데이터 표준요소(안)

앞 장에서 도출한 기반시설의 메타데이터 요소로는 명칭을 비롯하여 소재지 도로명주소, 위도, 경도, 데이터기준일자이고, 범주별 분석을 통하여 도출해낸 휴관일과 운영시간, 운영기관전화번호, 관리기관명칭 항목이 포함될 수 있다.

세부적으로 기반시설 관련 마스터데이터에서 활용할 수 있는 표준요소를 2장에서 정리한 마스터데이터의 특징과 비교하면 다음과 같다.

첫째, 명칭은 속성 값이 빈번하게 변경되는 항목이다. 기반시설의 경우 그 명칭이 사회적 분위기와 주변 의견에 따라 달라지기 쉽다. 이는 행정부처의 명칭 변화에서 쉽게 확인할 수 있다. 행정자치부는 1998년 내무부에서 행정자치부로 명칭이 변경되었고, 이후 행정안전부라는 명칭으로 개칭되었다. 2013년 3월 23일 안전행정부령 제1호 「안전행정부와 그 소속기관 직제 시행규칙」에 따라 행정안전부에서 안전행정부로 명칭의 변화가 있었다. 2014년 11월 19일 행정자치부령 제1호 「행정자치부와 그 소속기관 직제규칙」에 따라 안전행정부에서 행정자치부로 다시 명칭을 변경하였다. 상위 기관인 행정부처의 이름조차 빈번하게 변경되고 있는 상황에서 행정부처에 속해있는 기반시설의 명칭도 변할 가능성이 크다. 따라서 명칭을 마스터데이터 표준요소로 사용하기에는 적절하지 않다.

둘째, 소재지 도로명주소는 기본적으로 변경되기 힘든 속성을 지닌다. 도로명주소는 해당 지역의 다양한 부분을 고려하여 심의 후에 시·군·구청장이 새로운 주소를 부여하게 되어 있다. 「도로명주소법」에 따라 도로명주소의 변경을 위해서는 주소사용자 과반수의 서면 동의를 받아 도로명주소위원회의 심의를 거쳐야 한다. 다만, 도로명주소가 고지된 날부터 3년 이내에는 변경할 수 없다. 따라서 도로명주소를 변경하기 위해서는 많은 사람의 동의와 법적인 합의가 필요하며 많은 시간이 소요된다.

셋째, 위도 및 경도는 위치 속성을 의미한다

는 점에서 변경이 어렵다. 위도와 경도는 각도로 표현되며 지구상의 모든 장소는 위도와 경도를 이용해 표현할 수 있다. 국내에서만 사용되는 도로명주소와 달리 위도와 경도는 국제적으로도 통용되는 위치 표현법이며, 수치로 표현되는 것이기 때문에 위치 자체의 변경이외에는 수치 변화 경우는 존재하지 않는다.

넷째, 데이터기준일자는 데이터가 작성된 기준일자이다. 즉, 이 항목은 데이터 중 세부적인 항목이 변화할 때마다 빈번하게 변경되는 요소이다. 따라서 데이터기준일자 항목은 마스터데이터의 표준요소로 사용할 수 없다.

다섯째, 운영시간의 경우 이용자에게 서비스 업무상 빈번하게 사용되는 항목이지만 속성값이 쉽게 변하는 항목이다. 기반시설의 운영시간은 인원수의 감축이나 재정의 감축 등에 따라 변할 수 있는 속성이며 정형화된 데이터가 아니다. 따라서 운영시간 항목을 마스터데이터로 사용하기에는 한계가 존재한다. 다만, 공공기관의 경우 휴관일이 일정하게 정해져 있으므로 휴관일은 기본적으로 빈번하게 변경되지 않는 속성을 갖는다.

여섯째, 운영기관전화번호는 해당 기반시설을 대표하는 전화번호이다. 운영기관은 해당 기반시설을 실제로 운영하는 기관을 의미하며 운영기관의 전화번호는 기반시설에 관련된 정보를 바로 얻을 수 있는 전화번호를 말한다. 따라서 공공기관의 유선전화는 쉽게 변화하는 데이터가 아니며 전화번호 자체가 정형화되어 있는 숫자 데이터이다.

일곱째, 관리기관명칭에서 관리기관은 기반시설의 관리를 맡은 기관의 명칭을 말하는 것으로, 운영기관보다 상위의 전체적인 관리를

담당하는 기관이다. 구체적으로 지방자치단체나 행정부처 단위를 관리기관으로 지칭할 수 있다. 그러나 명칭은 일반적인 기반시설 명칭과 같이 쉽게 변화할 수 있는 항목으로 마스터데이터 표준요소로 활용하기에는 제약이 존재한다. 따라서 기관의 명칭과 관리기관에 관련된 정보는 쉽게 변화할 수 있는 명칭 자체를 사용하는 것이 아닌 행정표준코드로 대체하여 사용할 수 있다. 행정표준코드는 행정업무에 필요한 행정코드를 표준화한 정형화된 번호이다. 이 코드에는 기관코드가 포함되어 있는데 7자리의 숫자를 통해 기관을 표시한다.

표준 데이터세트만을 분석한 결과, 명칭과 데이터기준일자를 제외한 6개의 표준요소에 덧붙여 기반시설에 사용될 수 있는 마스터데이터 요소를 고민하여 보면 기반시설의 유형 항목이 추가될 수 있다. 기반시설 유형은 변경이 매우 어렵고 허용데이터를 제한함으로써 정형화된 데이터로서의 특성을 보일 수 있다. 기반시설 유형은 그 자체로 기반시설이 어떠한 목적을 지니고 있는지 완전하게 파악하게 해주는 속성이다. 기반시설 유형 정보를 표준 데이터세트에서 확인할 수 없던 이유는 이 유형정보가 데이터세트의 제목 및 항목명에 기입되어 있기 때문이다. 도서관의 경우 도서관이라는 구체적인 기반시설 유형 정보가 데이터세트의 제목인 '도서관', 항목명인 '도서관명'과 '도서관유형'에 포함되어 있다. 나머지 데이터세트에서도 마찬가지로 유형에 관련된 내용이 제목과 항목명에 기입되어 있으므로 개별적인 항목으로 찾아볼 수 없었다. 여기서 기반시설의 유형은 법령에서 명시하고 있는 유형으로 구체적으로 대통령령에서 세분화하고 있는 명칭을 사용할 수 있

다. 유형에 허용되는 데이터는 <표 2>의 대통령령 아래에서 확인할 수 있는 기반시설의 세부 항목으로 제한한다. 실제 분석한 표준 데이터 세트의 대다수가 대통령령에서 세분화하고 있는 분류체계에 속하고 있으므로 법률보다는 대통령령에 속한 분류가 활용도가 높다고 판단하였다.

종합분석결과와 범주분석결과, 이 연구에서는 기반시설의 메타데이터 표준요소를 기반시설 행정표준코드, 소재지도로명주소, 위도, 경도, 휴관 일자, 운영기관 전화번호, 관리기관 행정표준코드, 기반시설 유형의 총 8가지 항목으로 정리하였다. 8가지 항목의 표준요소를 개별적인 설명과 표현형식 및 단위와 종합하여 <표 8>로 정리할 수 있다.

건물이 있고 운영시간이 한정된 기반시설의 경우 휴관일과 운영시간, 운영기관전화번호, 관리기관명칭 항목이 추가될 수 있다. 휴관일 항목의 경우 휴관하는 날짜를 작성하거나 '연중무휴'라고 작성한다. 운영시간 및 이용시간의 경우 따로 항목을 제시하기보다는 자동차검사소 표준 데이터세트와 같이 하나의 값으로 표기한다. 운영기관 전화번호는 기반시설의 대표 전화번호이다. 관리기관 명칭은 전체 데이터세트를 분석하였을 때 관리기관이 존재하지 않는 전통시장을 제외하고 나머지 12개 데이터세트에서 확인할 수 있다.

본 논문의 경우 13개의 표준 데이터세트를 분석하여 추출한 마스터데이터 표준요소 수가 적게 구성되었다. 이는 분석대상과 관련하여 발생하게 되는 두 가지 한계점에 기인한 것이다. 첫째, 분석대상이 심층적이지 못하였다. 분석데이터로 제시한 표준 데이터세트는 이미 관련

〈표 8〉 기반시설 마스터데이터 표준요소(안)

항목명	설명	표현형식/단위
기반시설 행정표준코드	해당 기반시설의 행정표준코드, 행정표준코드관리시스템 참조	NNNNNNNN
소재지 도로명주소	소재지의 도로명과 건물번호로 표기한 주소	text
위도	해당기반시설 입구의 위도, WGS 84 위도좌표계 사용	NN.NNNNNN (소수점이하 6자리 이상 10자리 이하 허용)
경도	해당기반시설 입구의 경도, WGS 84 경도좌표계 사용	NN.NNNNNN (소수점이하 6자리 이상 10자리 이하 허용)
휴관 일자	해당기반시설이 휴관하는 날짜로 국가지정공휴일에만 휴관하는 경우 '국가지정공휴일'로, 휴관이 없는 경우 '연중무휴'로 표현, 공휴일 이외에 휴관하는 날짜는 "+"로 구분하여 표현	text
운영기관 전화번호	지역번호를 포함한 운영기관의 대표 전화번호	NNN-NNNN-NNNN
관리기관 행정표준코드	관리기관의 행정표준코드, 행정표준코드관리시스템 참조	NNNNNNNN
기반시설 유형	「국토계획법 시행령」에 따른 구분	text

된 데이터를 분석하고 공통으로 사용되는 데이터만을 추출한 것이다. 이러한 데이터를 대상으로 분석을 진행하였기 때문에 공통으로 나타나지 않은 세부적인 데이터 항목을 파악할 수 없다. 공통적인 데이터가 매우 적은 상태에서 새롭게 공통 항목을 뽑아냈기 때문에 나타날 수 있던 항목도 상대적으로 적다.

둘째, 분석대상이 광범위하다. 기반시설이라고 해도 세부적인 범주별로 공통 특징을 지닐 수 있다. 즉 본 논문에서는 표준 데이터세트와 기반시설로 분석대상을 한정하였기 때문에 더 많은 공통요소를 추출할 수 없는 한계를 갖는다. 기반시설 내에서 특정 시설 범주로 세분할 경우 공통으로 지니고 있는 항목을 더 많이 발견할 수 있다. 예를 들면 입장이라는 특성을 범주로 잡을 경우 입장료와 입장시간, 휴무일 항목이 공통 항목으로 추가될 수 있다. 이처럼 분석대상이 세부적이지 못하기 때문에 공통 항목이 많이 나타날 수 없었다.

4.2.2 기반시설 마스터데이터 표준요소 검토

위에서 제시한 표준요소가 범정부적으로 활용되기 위해서는 행정표준용어, 행정표준코드 및 공통항목 개방 기준과 같은 공공기관에서 활용하고 있는 표준화 체계를 준수하여야 한다. 개별적인 표준요소를 표준화 체계와 비교한 결과는 다음과 같다.

첫째, 기반시설 행정표준코드를 행정표준용어에서 살펴보면 기반시설과 행정표준코드를 찾을 수 있다. 행정표준용어에서 기반시설의 정의는 국토계획법의 법률 정의를 그대로 차용하고 있다. 행정표준코드의 정의도 「행정기관의 코드표준화 추진지침」의 법률 정의를 그대로 차용한다. 기반시설의 행정표준코드의 경우 이미 존재하는 행정표준코드를 활용하는 항목이고 공통항목 개방 기준에서 정의하고 있지 않은 항목이기 때문에 세 가지 표준화 체계에 부합한다.

둘째, 소재지 도로명주소를 행정표준용어와

비교하면 소재지와 도로명주소로 구분하여 확인할 수 있다. 소재지는 주요 건물이나 기관 따위가 자리 잡고 있는 곳으로, 도로명주소는 도로명 사업에 의하여 새로이 부여된 건물번호에 의하여 표기하는 주소로 정의한다. 행정표준코드는 사용하지 않으며 공통항목 개방기준 중 11번째의 도로명주소 개방항목의 설명에 따라 정의한다.

셋째, 위도를 행정표준용어와 대조해본 결과 표준어로 정의되고 있다. 위도는 공통항목 개방 기준에서 정의하고 있지 않는 항목이기 때문에 모든 데이터셋에서 공통으로 정의하는 설명을 차용하였다.

넷째, 경도도 위도와 마찬가지로 행정표준용어와 대조해본 결과 표준어로 정의하고 있다. 또한 공통 항목 개방 기준에서 정의되고 있지 않는 항목이기 때문에 위도와 같이 모든 데이터셋에서 공통으로 정의하는 설명을 차용하였다.

다섯째, 휴관 일자 항목은 행정표준용어에서 휴관과 일자로 나누어 찾을 수 있다. 휴관은 도서관이나 영화관 따위에서 하던 일을 하루나 한동안 쉬므로 정의하고 있고, 일자는 연월일로 정의하고 있다. 휴관 일자의 경우 공통항목 개방 기준 중 세 번째에 정의된 일자의 기준을 따른다. 일자는 일정한 기간으로서의 어느 한 해와 달의 하루를 의미하며 YYYY-MM-DD로 표기할 수 있다. 휴관 일자는 휴관하지 않는 기반시설, 국가지정공휴일에만 쉬는 기반시설, 국가지정공휴일 이외에 정기적인 휴일이 있는 기반시설, 정기적이지 않은 휴일이 있는 기반시설 등에 따라 자유롭게 기재할 수 있게 정의하였다. 각 휴관일은 '연중무휴'와 '국가지정공

휴일' 등의 텍스트를 덧셈(+) 표시를 이용하여 구분할 수 있다.

여섯째, 운영기관 전화번호는 운영기관의 대표 전화번호를 기재하게 되며, 전화번호는 공통항목 개방 기준 중 15번째의 전화번호 정의에 따른다. 개방기준에서는 통신망을 통한 상호접속을 위하여 이용자에게 부여된 식별번호로 전화번호를 정의한다. 행정표준용어에서 운영, 기관, 전화번호로 나누어 살펴볼 수 있다.

일곱째, 관리기관 행정표준코드의 경우 기반시설 행정표준코드와 마찬가지로 행정표준코드를 준수하여 기재하는 항목이다. 행정표준용어에서 살펴보면 관리와 기관으로 나누어 확인할 수 있다.

여덟째, 기반시설 유형 항목은 「국토계획법 시행령」에 따라 41개의 항목 중 선택하여 기재한다. 이는 행정표준코드와 공통항목 개방 기준과 관련되어 있지 않은 항목이며, 행정표준용어와 비교하면 기반시설과 유형으로 나누어 확인할 수 있다.

본 논문에서 제안하고 있는 8가지 항목은 모두 행정표준용어, 행정표준코드, 공통항목 개방 기준의 표준화 체계를 준수하고 있는 항목이다. 따라서 이 항목은 범정부적으로 이미 존재하고 있는 데이터를 종합할 때 많은 변화 없이 정리할 수 있고, 기존의 표준화 체계를 준수하고 있다는 점에서 대표성을 지닐 수 있다.

5. 결 론

이 연구에서는 기반시설 마스터데이터 구축을 위한 프로세스를 제공하고 기반시설 마스터

데이터의 구축 가능성을 검증하였다. 마스터데이터는 여러 업무에서 공통으로 참조될 수 있는 중요 데이터로 쉽게 변화하지 않는 특성을 보이는 데이터이며, 전자형태로 이루어진 정부 생산·관리 데이터를 공유하여 활용할 수 있는 기준데이터이다. 또한 마스터데이터는 범정부 차원에서 전반적으로 빈번하게 사용되는 데이터라는 특징을 지닌다. 기반시설은 법령으로 규정된 항목이라는 점에서 객관적인 범주로 작용할 수 있으며 변화 가능성이 적고, 개별적인 기반시설의 운영·관리 기관이 다양하므로 관련 데이터가 여러 기관에 분산·관리되고 있다. 따라서 이 연구에서는 기반시설을 대상으로 마스터데이터 구축 시에 고려할 항목과 요소를 도출하였으며 연구결과를 요약하면 다음과 같다.

마스터데이터 구축을 위하여 우선 13개의 기반시설 관련 표준 데이터셋을 분석하고, 이를 마스터데이터의 특징과 비교하여 기반시설 마스터데이터의 표준요소로 활용할 수 있는 항목을 선정하였다. 그 결과 기반시설 행정표준코드, 소재지 도로명주소, 위도, 경도, 휴관 일자, 운영기관 전화번호, 관리기관 행정표준코드의 항목을 기반시설 마스터데이터 표준요소로 도출하였다. 또한 필수 항목으로 기반시설 유

형 항목도 제안하였다. 본 논문에서 제시한 8가지 항목은 공공기관의 표준화 체계에 부합한다는 점에서 많은 비용이나 시간을 소요하여 기존의 데이터에 수정을 기할 필요 없이 사용할 수 있는 장점을 지닌다.

다만 본 논문에서 추출한 표준요소는 심층적이지 못하고 광범위한 분석대상으로 인하여 적은 요소만 추출되었다는 제한점을 지니고 있다. 이러한 제한점에도 불구하고 본 논문에서는 마스터데이터 표준요소로 활용할 수 있는 항목을 제안하고 있다는 점에서 의미를 찾을 수 있다. 마스터데이터 표준요소 항목은 기본적인 시설 항목을 대상으로 한 마스터데이터에서 활용할 수 있으며, 추후 연구를 진행하는 데 기본적인 항목으로 작용할 수 있다. 본 논문의 제한점에 기반을 두어 이후 논문에서 이루어져야 하는 연구를 제안하자면 첫째, 국가 마스터데이터 구축을 위하여 전반적인 공공기관 업무 프로세스 분석과 함께 구체적인 선정 기준에 관한 규정이 이루어져야 한다. 둘째, 특정 범주의 마스터데이터 구축 시 공공데이터포털, 정보공개포털, 개인정보보호종합포털, EA포털, 행정정보공동이용센터 등에 등록된 데이터 현황을 종합적으로 조사하여 분석하는 연구의 필요성이 존재한다.

참 고 문 헌

공공데이터포털 [online]. [cited 2017.5.1.] <<http://data.go.kr>>.
 박진경, 김미수. 2016. 조립형 ERP를 위한 마스터데이터 관리 API. 『한국컴퓨터정보학회논문지』, 21(11): 93-101.

- 손혜인. 2017. 『기반시설 마스터데이터 표준요소 구축에 관한 연구』. 석사학위논문. 중앙대학교 일반대학원, 문헌정보학과.
- 이승민. 2013. 마스터데이터를 활용한 메타데이터 통합 프레임워크 구축. 『한국도서관·정보학회지』, 44(1): 201-225.
- 조현구. 2006. 마스터데이터 통합 솔루션. [online]. [cited 2017.5.10].
〈[www.kmis.or.kr/3_sig/erp_data/erp\(jun06\)2.pdf](http://www.kmis.or.kr/3_sig/erp_data/erp(jun06)2.pdf)〉.
- 최영훈, 김태진, 서승현, 정세희. 2015. 『데이터 관리체계 개선방안 연구』. 행정자치부, 연구보고서 11-1740000-000188-01.
- 한국기록학회. 2008. 『기록학 용어 사전』. 서울: 역사비평사.
- 행정자치부 [online]. [cited 2017.5.1]. 〈<http://moi.go.kr>〉.
- 행정표준코드관리시스템 [online]. [cited 2017.5.10]. 〈<https://www.code.go.kr>〉.
- Berson, A. and L. Dubov. 2007. *Master Data management and customer data integration for a global enterprise*. New York: McGraw-Hill.
- Cervo, D. and M. Allen. 2011. *Master Data management in practice: Achieving true customer MDM*. NJ: John Wiley and Sons.
- Dreibelbis, A., E. Hechler, I. Milman, M. Oberhofer, V. P. Run, and D. Wolfson. 2008. *Enterprise Master Data Management*. IN: IBM Press.
- Haug, A. and J. S. Arlbjorn. 2011. "Barriers to master data quality." *Journal of Enterprise Information Management*, 24(3): 288-303.
- Loshin, D. 2009. *Master Data management*. Morgan Kaufmann.
- Ofner, M., K. Straub, B. Otto, and H. Oesterle. 2013. "Management of the Master Data Life Cycle: a Framework for Analysis." *Journal of Enterprise information Management*, 26(4): 472-491.
- Oracle. 2011. Master Data management. [online]. [cited 2017.4.22].
〈<http://www.oracle.com/us/products/applications/master-data-management/018876.pdf>〉.
- Rivas, B., J. Merino, I. Caballero, M. Serrano, and M. Piattini. 2016. "Towards a service architecture for Master Data exchange based on ISO 8000 with support to process large datasets." *omputer Standards & Interfaces*, 54: 94-104.
- Silvola, R., O. Jaaskelainen, K. Kropsu-Vehkapera, and H. Haapasalo. 2011. "Managing one master data - challenges and preconditions." *Industrial Management and Data Systems*, 111(1): 146-162.
- Vilminko-Heikkinen, R. and S. Pekkola. 2017. "Master Data management and its organizational implementation." *Journal of Enterprise Information Management*, 30(3): 454-475.

Wang, L., X. Ming, and J. You. 2009. "The steps and methodology of identifying Master Data from business processes." *World Congress on Software Engineering*, 329-333.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

Administrative standard code management system [online]. [cited 2017.5.10].

〈<https://www.code.go.kr>〉.

Cho, Hyeon-Gu. 2006. Master data consolidation solution. [online]. [cited 2017.5.10].

〈[www.kmis.or.kr/3_sig/erp_data/erp\(jun06\)2.pdf](http://www.kmis.or.kr/3_sig/erp_data/erp(jun06)2.pdf)〉.

Choe, Yeong-hun, Taejin Kim, Seunghyeon Seo, and Sehui Jung. 2015. *A Study on the Improvement of Data Management System*. Department of Government Administration and Home Affairs, Research Report 11-1740000-000188-01.

Korean Society Of Archival Studies. 2008. *Records and Archives Management Glossary*. Seoul: Yeoksabipyongsas.

Lee, Seungmin. 2013. "Construction of Framework for Metadata Integration Using Master Data Approach." *Journal of Korean Library and Information Science Society*, 44(1): 201-225.

Ministry of Government Administration and Home Affairs [online]. [cited 2017.5.1].

〈<http://moi.go.kr>〉.

Park, Jin-Kyung and Mi-su Gim. 2016. "Master Data Management API for Assembled ERP Tools." *Journal of The Korea Society of Computer and Information*, 21(11): 93-101.

Public Data Portal [online]. [cited 2017.5.1]. 〈<http://data.go.kr>〉.

Sohn, Hye-In. 2017. *A study on the Establishment of Standard Elements of Infrastructure Master Data*. Master diss., Chung-Ang University Graduate School, Library and Information Science Department.