

데이터 유통 플랫폼을 위한 객체 식별자

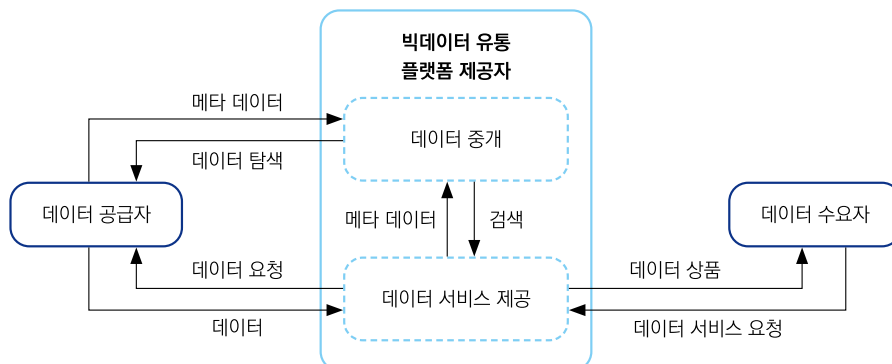
유시형 한국지능정보사회진흥원 지능데이터기반팀 수석연구원

1. 머리말

빅데이터 유통 생태계는 빅데이터 상품을 제공하고 소비하는 과정에서 구성원이 수행하는 역할의 유기적 관계에 따라 구성된다. 이에 따라 빅데이터 유통 생태계의 구성 요소는 데이터 공급자, 빅데이터 유통 플랫폼 제공자 및 데이터 수요자로 구분할 수 있다. 데이터 공급자는 빅데이터 유통 플랫폼 제공자의 데이터 중개 부문이 보유한 메타데이터를 공유한다. 데이터 중개 부문은 해당 메타데이터를 기반으로 데이터 공급자의 데이터를 탐색한다. 데이터 공급자는 데이

터 서비스 제공 부문에 데이터를 제공하며, 데이터 서비스 제공 부문은 데이터 수요자로부터 데이터 상품 요청에 따라 데이터를 제공한다. 데이터 수요자는 데이터 상품 서비스 부문에 데이터 상품을 요청하며 데이터 서비스 제공 부문은 이에 따라 수요자에게 데이터 상품을 공급한다. [그림 1]은 빅데이터 유통 생태계의 일반적인 구조를 보여준다.

다수의 데이터 유통 플랫폼 제공자가 시장에 존재해서 다양한 데이터 공급자 및 수요자와 거래할 경우, 제공되는 데이터에 대한 거버넌스 체계가 필요하다. 일관된 데이터 거버넌스 체계가



[그림 1] 빅데이터 유통 생태계 모델(예)

<표 1> 객체식별자(OID) 국내 할당 현황

번호	OID	용도	관리기관
1	{0.2.450}	국내 RFID 코드 식별	한국인터넷진흥원
2	{0.2.480}	정보통신단체표준 및 기술보고서 식별	
3	{0.2.481}	사물인터넷 디바이스 식별	
4	{2.27}	국제 RFID 하위 코드 식별	
5	{1.2.410}	전자상거래, 보안 및 통신망 관리 식별	국가기술표준원
6	{0.3.4500}	KT 식별	네트워크 사업자
7	{0.3.4501}	DACOM 식별	
8	{0.3.4501}	CSDN(circuit switched data network) 식별	
9	{2.49.0.0.410}	기상경고 식별	기상청
10	{2.16.410}	미정(미사용)	미부여

존재하지 않을 경우 복수의 데이터가 유통될 때 동일하거나 유사한 데이터가 시장에 공급되어 자원 배분의 효율성을 떨어뜨릴 수 있기 때문이다. 이는 수요자가 적절한 데이터를 탐색하고 이용하는 데 저해 요인으로 작용할 수 있다. 따라서 전체 데이터를 일관적으로 식별할 수 있는 데이터 식별자 체계는 빅데이터 유통 생태계의 필수적인 요소라고 할 수 있다.

2. 객체식별자(OID) 적용 데이터 식별자

2020년 7월 14일 발표된 한국판 뉴딜 사업 중 ‘데이터 댐’ 과제는 2025년까지 분야별 빅데이터 플랫폼을 10개에서 30개로 확대하고, 공공데이터 14.2만 개를 조기에 개방하며, AI 학습용 데이터 1,300종을 구축함으로써 분산된 데이터 연계·검색을 할 수 있는 통합 플랫폼을 구현하는 것이 골자다.

공공부문을 필두로 다수의 데이터 플랫폼과 데이터 포털이 등장하면서 유사하거나 중복되는 데이터 생산 및 유통 문제가 대두되기 시작했다. 현재 ‘공공데이터의 제공 및 이용 활성화에 관한 법률’ 제15조의2(중복·유사 서비스 개발·제공의

방지), 제15조의3(중복·유사 서비스의 개발·제공에 관한 실태조사 등)에서는 중복·유사 서비스 조사 및 정비를 위한 빅데이터와 공공데이터, 민간데이터 간 주기적 실태조사, 온라인 창구를 통한 상시신고 등을 이에 대한 대책으로 제시된다.

그러나 데이터 댐 사업으로 다양한 분야로부터 생성될 수많은 데이터의 중복, 유사 데이터에 대한 식별체계를 구체적으로 정의하지 않은 채 데이터명, 혹은 데이터설명 정도에만 의존하여 데이터를 관리한다면 한계에 봉착할 수밖에 없다. 현 상황에서는 중복·유사 데이터가 생성되지 못하도록 제한하는 것이 사실상 불가능하므로 데이터 댐 과제 추진 과정에서 문제점이 지속적으로 야기되어 사업의 성과가 전체적으로 저하될 수 있을 것이다. 이에 따라, 중복·유사 데이터가 생성되지 않도록 기존 방법과 차별화된 고유한 식별체계인 객체식별자(OID, Object Identifier) 차원의 접근이 필요하다. 중복·유사 데이터 생산 방지 및 데이터 관리·연계·활용을 위한 데이터 식별체계 구축 방안으로 객체식별자(OID) 적용을 고려하는 것이다.

객체식별자는 ‘지속성’ 있는 이름이 필요한 전 세계적으로 ‘모호하지 않은’ 이름을 가진 ‘객체’,

<표 2> 객체 식별자 {2 16 410 n} 할당 예시

1차 정수값 (Primary Integer value)	2차 식별자 (Secondary identifier)	설명
1	public	국가 및 공공
2	private	민간 기관 및 기업
3	individual	개인

‘개념’ 또는 ‘사물’의 모든 유형에 이름을 부여하기 위해 ITU-T와 ISO/IEC가 공동으로 개발한 고유한 식별 메커니즘이다. 객체식별자 국내 적용 현황은 국제적으로 할당받은 국내 고유 식별 체계 총 10개 중 <표 1>과 같이 활용되고 있으며 {2.16.410}는 국내에 할당되어 있으나, 사용되지 않는 것으로 확인된다. 따라서 유희 자원인 {2.16.410}을 데이터 식별체계에 적용할 수 있다.

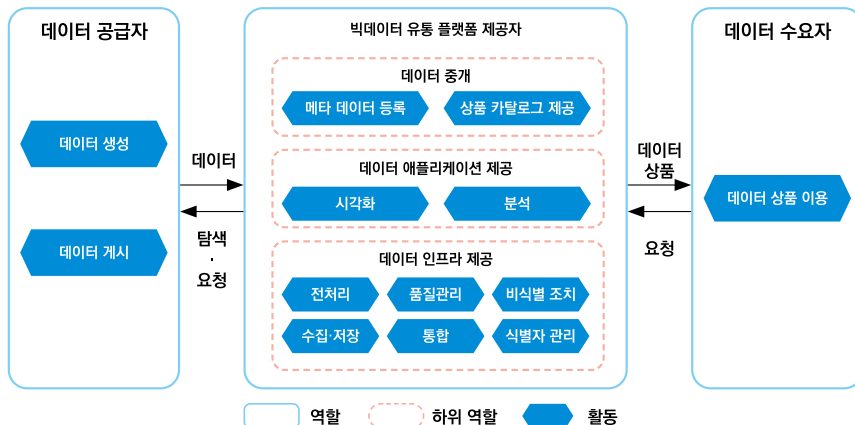
3. 빅데이터 유통 플랫폼에서 데이터 식별자 역할

서두에서 언급했듯이 빅데이터 유통 생태계는 데이터 공급자, 빅데이터 유통 플랫폼 제공자 및 데이터 수요자로 구성된다. 빅데이터 유통 플랫폼 제공자의 하위 역할인 데이터 중개, 데이터 애플리케이션 및 데이터 인프라는 빅데이터 유통 플랫폼 서비스를 이용하는 데이터 공급자 및

수요자를 대상으로 제공된다. 빅데이터 유통 플랫폼 제공자의 하위 역할인 데이터 인프라 제공의 ‘식별자 관리’ 부문은 객체식별자 체계와 연동된다.

빅데이터 유통 플랫폼은 데이터 수집, 생산, 유통 및 관리를 위한 객체식별자를 부여하고 이를 해석시스템을 통해 관리한다. 데이터 유통을 위한 객체식별자는 데이터 생산 및 융합 주체 기준으로 의미적 분류체계에 따라 부여되며 데이터의 전반적 현황에 대한 정보를 제공한다. 객체식별자를 관리하려면 데이터 식별을 위한 객체식별자 체계, 데이터 식별을 위한 객체식별자 관리절차 및 데이터 식별을 위한 이력관리 항목과 같은 정보가 필요하다.

4. 객체식별자(OID) 할당



[그림 2] 빅데이터 유통 생태계 구성원 역할 및 활동

<표 3> 객체식별자 {2 16 410 1 n} 할당 예시

하위 아크	1차 정수값 (Primary Integer value)	2차 식별자 (Secondary identifier)		변환 규칙
		기관명	행정표준코드	
n	1	대통령	0000001	앞자리 0 전부 삭제
	11000000	자치정보화조합	A000000	A를 11로 교체
	12000000	정부산하기관위원회	B000000	B를 12로 교체
	13000000	부경대학교 누리해양바이오식의약사업단	C000000	C를 13로 교체
	14000000	국민건강보험공단 본부	D000000	D를 14로 교체
	15000000	우리은행	P000001	E를 15로 교체
	1600189	감찰위원회	1Z00189	1Z를 16로 교체

<표 4> 객체식별자 {2 16 410 11 n m l} 할당

하위 아크	1차 정수값 (Primary Integer value)				2차 식별자 (Secondary identifier)
n	데이터 분류체계 예시				N/A
	정치 : 1	농림수산 : 11	환경 : 21	문화 : 31	
	행정 : 2	자원에너지 : 12	가정세대 : 22	종교 : 32	
	사법 : 3	기계 : 13	의(依) : 23	교육 : 33	
	외교 : 4	운수교통 : 14	식(食) : 24	학술 : 34	
	군사 : 5	토목건설 : 15	주(住) : 25	예술 : 35	
	경제 : 6	상업 : 16	생활용품 : 26	스포츠(운동) : 36	
	금융 : 7	정보산업 : 17	레저(여가) : 27	자연 : 37	
	기업 : 8	사회 : 18	행사 : 28	생물 : 38	
	노동 : 9	건강 : 19	범죄 : 29	국제 : 39	
산업 : 10	질병 : 20	재해 : 30	국내지역 : 40	국외지역 : 41	
m	데이터 등록 일련번호 예. 1234				N/A
l	계정번호 예. 1				계정번호 예. r1

데이터 세트 단위로 부여되는 객체식별자는 하위에 국가 기관, 민간단체 및 조직 그리고 개인을 대상으로 ‘아크’를 할당할 수 있다. 본 표준에서 정의하는 객체식별자 {2 16 410}의 하위 아크는 <표 2>와 같다

객체 식별자 {2 16 410 1 n}은 데이터를 생산하고 등록하는 국가 및 공공기관을 식별하기 위해 사용한다. 국가 및 공공기관 번호 부여 체계는 국가 및 공공기관을 아우르는 각급 기관의 행정코드를 표준화하여 제정, 고시한 행정표준코드(행정표준코드관리시스템, <https://www.code.go.kr/index.do>)를 준용한다. 다만, 앞자리가 0으로 시작하는 국가기관 코드와 알파벳

으로 표시된 공공기관 코드는 변환하여 적용한다. 앞자리가 0으로 시작하거나 정수가 아닌 알파벳을 포함하는 경우, 앞자리에 위치한 0을 전부 삭제하거나 알파벳을 숫자로 변환하여 적용한다.

대통령이 생산·등록한 데이터의 객체 식별자 {2 16 410 1 1 n} 하위 아크는 생산한 데이터의 의미적 분류체계에 따라 부여되고, 이후 해당 분류체계 내의 일련번호가 할당된다. 데이터 분류체계는 국가기록원의 일반주제어 시소러스의 일반어 대상 주제분류를 준용한다. 만약 등록된 해당 데이터가 개정될 경우 개정번호가 부여된다.

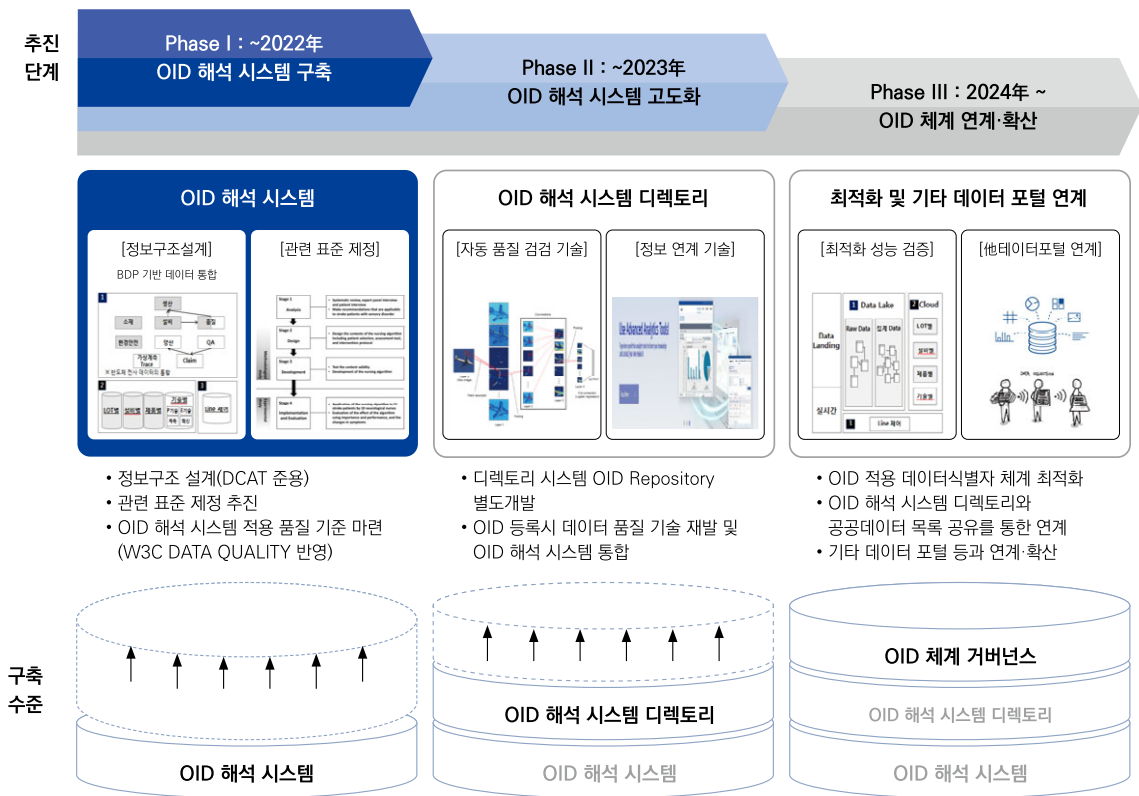
자치정보화조합이 행정 관련 데이터를 생산하

여 행정 분류체계에서 3번째로 등록한다면 {2 16 410 1 11000000 2 3}으로 식별되고, 해당 데이터를 1차 개정한다면 {2 16 410 1 11000000 2 3 1}로 식별된다. 민간 및 기업 대상 객체 식별자 {2 16 410 2 n}은 민간기관 및 기업이 생산·등록하는 데이터를 식별하기 위해 사용한다. 등록 순서대로 기관 식별을 위한 객체식별자 {2 16 410 2 n}을 발급하고 하위에서는 기관 구분의 일련번호를 부여하여 이후 <표 4>와 같은 하위 아크를 할당할 수 있다. 객체식별자 {2 16 410 3 n}은 개인이 생산·등록하는 데이터를 식별하기 위해 사용한다. 등록 순서대로 개인식별을 위한 객체식별자 {2 16 410 3 n}을 발급하고 하위에서는 개인 구분의 일련번호를 부여하여 이후 <표

4>와 같은 하위 아크를 할당할 수 있다.


5. 객체식별자 해석시스템(ORS : Object identifier Resolution System) 구축

할당된 객체식별자를 빅데이터 생태계에 적용하려면 객체식별자 해석시스템(ORS)이 필수적이다. 중앙집중방식으로 데이터 식별체계를 부여하고 관리하는 객체식별자 해석시스템은 국가 차원의 데이터 관리 기반과 데이터 정책 수립에 근거를 제공할 수 있다. 객체식별자 해석시스템 구축 로드맵은 크게 시스템 구축, 시스템 고도화, 체계 연계 및 확산의 3단계로 구성된다. 시스템 구축 단계는 DCAT(Data Catalog Vocabulary)을 반영하여 시스템 정보 구조 설계, 관



[그림 3] 객체식별자 해석시스템 구축(안)

런 표 제정 및 해석시스템 등록 데이터 품질 기준 마련 등의 세부과제로 구성된다. 시스템 고도화 단계는 객체식별자 해석시스템 디렉토리를 개발하고, 자동화된 데이터 품질 점검 기술을 개발한 후 해석시스템에 통합하는 작업을 포함한다. 최종 단계인 해석시스템 최적화 및 연계·확산 단계는 객체식별자 해석시스템 최적화, 해석시스템 디렉토리 및 공공데이터 포털 데이터목록 공유를 통한 연계, 공공·민간을 아우르는 데이터 포털과 연계 및 객체식별자 체계 확산 등으로 구성된다.

객체식별자 해석시스템 등이 요건하고 필수적인 국가자원으로 활용되기를 기대해본다. 

6. 맺음말

객체식별자 해석시스템을 포함한 데이터 유통 플랫폼 구축을 위해 2019년 10개 빅데이터 플랫폼을 필두로 향후 30개의 플랫폼을 구축하여 데이터 수집·유통·활용 기반 조성 및 데이터 표준화를 추진한다. 공공분야뿐 아니라 대기업을 비롯한 민간분야에서도 데이터 포털과 같은 데이터 유통플랫폼을 선보이고 있다. 이러한 상황에서 빅데이터 플랫폼에 적용될 수 있는 객체식별자 체계 및 객체식별자 해석시스템은 중앙집중방식으로 데이터 인프라를 구축하고 관리하는 데 근거로 활용될 수 있다. 이를 통해 향후 펼쳐질 데이터 중심 경제에서 객체식별자 체계와

참고문헌

- [1] ITU-T X.660 | ISO/IEC 9834-1:2012, 'Information technology - Procedures for the operation of object identifier registration authorities: General procedures and top arcs of the international object identifier tree', 2011.
- [2] ITU-T, Object identifiers (OIDs) and their registration authorities, ITU-T Handbook, 2010.
- [3] 백형중·정유경, 산업간 융합을 위한 식별체계인 OID 이해 및 발전방향, Internet & Security, 2013.
- [4] ITU-T X.672 | ISO/IEC 29168-1: 2011, 'Information technology - Open systems interconnection - Part 1: Object identifier resolution system', 2011.