

# 부록 이용자중심 메타데이터

### 데이터 거래 지원 가이드라인





※ 본 가이드는 데이터 유통 및 활용 촉진을 지원하기 위한 연구 결과물로서, 데이터 거래 시 참고할 수 있도록 가격·품질·법적 쟁점 등의 이해를 돕기 위한 자료입니다.



부록. 이용자중심 메타데이터



# CONTENTS

제 1장	추진배경 및 목적	
	1.1. 추진 배경 1.2. 추진 목적	1 2
제 2장	이용자 중심 메타데이터	
	2.1. 개요2.2. 적용 대상2.3. 적용 범위2.4. 가이드 특징	3 4 4 4
제 3장	현황 및 검토	
제 3장	현황 및 검토  3.1. 국내 거래소 현황  3.2. 데이터 카탈로그  3.3. 데이터 카탈로그와 메타데이터  3.4. 메타데이터의 필요 정보	5 6 7 8
제 3장 제 4장	3.1. 국내 거래소 현황 3.2. 데이터 카탈로그 3.3. 데이터 카탈로그와 메타데이터	6 7





Korea Data Agency

## 추진배경 및 목적

#### 1.1. 추진 배경

메타데이터는 데이터관리의 핵심 요소로써 초기 메타데이터는 데이터의 구조와 형식을 관리하는 기술(테크니컬) 메타데이터로 출발하였으며 이는 기업의 물리적인 시스템 내 데이터 관리를 위한 필수적인 요소가 되었다.

점차 데이터의 시스템적인 관리 측면 보다 데이터의 이용과 활용에 대한 요구가 커짐에 따라 기술적인 기반이 없는 현업 사용자들을 위해 공통의 의사소통을 위한 정의가 필요해졌다. 그러한 맥락에서 물리적인 데이터에 대한 손쉬운 이해와 접근을 위해 업무적인 용어를 사용하여 정의하고 표준화하여 활용하는 것이 보편화 되었으며 이는 비즈니스 메타데이터로 확장하며 Self-BI의 핵심 요소로 자리 잡았다.

하나의 조직 내 구성원들은 서로 같은 업무 환경하에 해당 기업이 속한 업종의 고유한 용어들을 공유하기 때문에 비즈니스 메타데이터 관리에서 업무용어를 표준화하고 활용하는 것이 비교적 수월하다 할 수 있다.

아울러, 동일 업종의 기업들은 거의 동일한 용어를 사용하여 업무를 수행하기 때문에 업종 내 기업 간 데이터를 이해하는 것 또한 비교적 수월하다 할 수 있다.

하지만 오늘날의 비즈니스 환경은 업종의 경계가 희미해지고 서로 다른 업종 간의 데이터 를 융·복합 분석하여 새로운 비즈니스 기회를 창출하고자 하는 요구들이 지속해서 증가하고 있다.

이를 위한 데이터의 안정적인 유통 기반을 마련하고자 데이터거래소 및 빅데이터 플랫폼이 출현하고 있으며 이에 맞춰 거래소 간 데이터의 상호연계를 위한 표준으로 DCAT(Data Catalog Vocabulary)이 자리 갑게 되었으며 국내 빅데이터 플랫폼을 비롯한 데이터거래소 들 또한 이러한 흐름에 따라 메타데이터 관리 및 유통을 위한 카탈로그 시스템에 DCAT 표준을 적용하는 등 대응해 나가고 있다.

데이터 거래의 관점에서 보면 데이터 상품이라는 특성에 따른 차이점은 있겠지만 궁극적인 "거래"의 개념은 실물상품의 "거래"와 별반 다르지 않다.

데이터 상품을 공급하는 공급자가 있고, 이용 수요를 갖고 있는 구매자가 있으며, 그러한 거래를 중개하는 시장(마켓플레이스)으로서의 거래소가 존재한다.

대부분 실물경제에서의 거래행태나 트렌드를 볼 때 좋은 상품의 공급이 곧바로 수요로 이어지던 공급자 주도 시장에서 구매자의 수요와 욕구를 분석하여 그에 부합하는 상품을 만들기 위해 노력해야 하는 구매자 주도 시장으로의 전환은 이미 오래전에 도래하였고 더욱 가속화되고 있다.

데이터경제를 표방하는 현시점에서 이러한 실물경제에서의 트렌드가 데이터 거래에서도 피할 수 없는 현실임은 자명하다.

이러한 배경 아래에서 데이터거래를 촉진하기 위해서는 구매자의 분석 수요에 부합하는 데이터가 제공되어야 하며, 구매자가 수요에 부합하는 데이터를 찾고 이용함에 있어 필요한 정보를 명확하고 적절하게 전달하는 것이 메타데이터의 역할이라 할 수 있다.

#### 1.2. 목적

본 안내서는 구매자가 적합한 데이터 상품을 찾고, 이해하고, 활용하기 위한 메타데이터 항목을 정의하고 공급자 측면에서 이러한 정보를 충실히 제공하도록 가이드라인을 제시하고 자 한다.

- 구매자가 데이터 상품 구매 시 목적 적합성 및 구매가치를 판단할 수 있는 정보를 제공 한다.
- 공급자가 데이터 상품에 대한 정보제공 시 필요항목 및 수준에 대한 기준을 제공한다.





Korea Data Agency

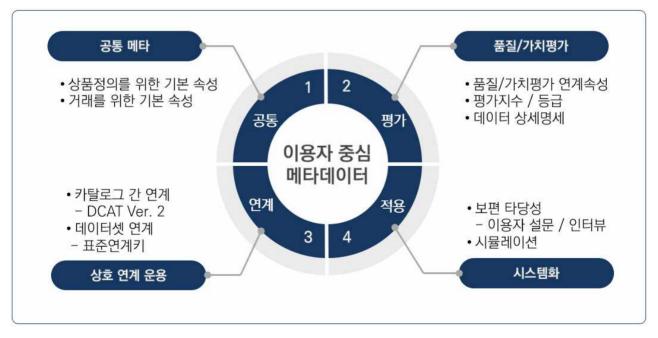
### 이용자 중심 메타데이터

#### 2.1. 개요

메타데이터는 데이터관리 및 검색, 평가, 이해를 위한 데이터의 모음으로 데이터 이용자가데이터를 검색하고 의도된 용도에 맞는 데이터의 적합성을 평가하기 위한 정보를 제공한다.

메타데이터의 사전적인 의미를 떠나 본 안내서에서의 메타데이터는 데이터 상품 거래를 위해 이용자에게 제공되어야 하는 데이터를 통칭한다.

메타데이터는 데이터 상품 정의와 거래를 위한 기본적인 공통 속성을 관리하여야 하고, 데이터 상품에 대한 목적 적합성을 평가하기 위한 속성을 가져야 하며, 데이터 간 연계를 위한 확장을 고려해야 하고, 실제 시스템으로 적용이 가능한 보편타당한 구성이어야 한다.



[그림-부록-1] 이용자 중심 메타데이터

#### 2.2. 적용 대상

- 메타데이터는 데이터상품 거래 및 유통을 위한 거래소(유통플랫폼)의 데이터 상품 관리를 위한 체계로 활용한다.
- 메타데이터는 공급자의 상품등록을 위한 프로세스의 출발점이 되며 주어진 정보를 기반으로 품질/가치평가를 연계한다.
- 메타데이터는 상품 거래를 위한 기본(필수) 정의 항목 외에도 상호 연계 운용을 위한 항목, 품질/가치평가를 위한 상세명세 속성 등으로 구성되어 있어 상황에 따라 항목을 선별 적용할 수 있으나 구매자 관점에서 데이터 상품에 대한 명확한 이해와 활용성을 판단할 수 있도록 최대한 명확히 등록할 수 있도록 하여야 한다.

#### 2.3. 적용 범위

- 데이터 이용자가 데이터를 찿고, 이해하고, 평가할 수 있는 수단으로서의 메타데이터 체
   계
- 데이터 준비 및 데이터 융합을 위한 유용한 정보를 공유하기 위한 리소스
- 데이터 유통 생태계의 이해관계자 간 효과적인 커뮤니케이션을 보장하는 표준 및 가이 드라인

#### 2.4. 가이드 특징

- 데이터 거래소(유통플랫폼) 측면에서 중개 데이터 상품에 대한 공급자 상품등록, 품질/ 가치평가 모형 연계, 상품 검색 및 관리 기능 구현 시 기준 메타데이터로 활용할 수 있 다.
- 공급자 측면에서 데이터 상품에 대한 품질관리, 접근성 관리 등 시장가치를 높이기 위해 제공해야 할 메타데이터의 항목 및 수준에 대한 가이드로 활용할 수 있다.
- 구매자 측면에서 데이터 상품에 대한 명확한 정보를 수집하고, 활용성을 판단하고, 구매 가치를 판단하기 위한 신뢰 할 수 있는 기준 정보로 활용할 수 있다.



PART 03

Korea Data Agency

# 현황 및 검토

#### 3.1. 국내 거래소 현황

국내 현황분석 대상 11개 데이터거래소 및 플랫폼에 대한 데이터 카탈로그의 항목을 분석한 결과 기본적으로 데이터 상품의 판매 및 유통을 위한 메타데이터를 관리하고 있음. 또한, 선행 연구를 통해 데이터 간 상호 공유 및 교환을 목표로 DCAT 표준을 규격화한 메타데이터를 정의하고 빅데이터플랫폼 등에 적용하고 있는 상황임.

다만, 데이터의 관리와 상호운용성에 대한 구조적 기반을 마련한 것은 성과로 나타나고 있으나 이용자의 데이터 활용 측면에서는 다소 미비한 측면이 있다.

이용자의 활용 혹은 거래 활성화 측면을 고려할 때 국내 거래소 및 유통플랫폼의 메타데이터가 공급자 중심의 정보제공에 치우쳐 있으며 데이터 상품의 품질 및 가치 판단을 위한 세부명세를 관리하지 않거나 제공하지 않는 것으로 나타났다.

데이터셋 자체에 대한 명확한 이해와 활용 적합성을 판단하기 위한 정보제공 면에서는 보완이 필요하며, 선행 연구들에서 거래 활성화의 저해 요인으로 꼽고 있는 품질/가격의 불합리성에 대한 우려를 해소하기 위한 정보제공 측면에서도 다소 부족하다.

본 연구에서는 이러한 점을 보완하고자 구매자가 데이터 상품을 이해하고 활용하기 위한 ① 구매자 중심의 메타데이터 항목을 추가 도출하였으며 ② 데이터 상품의 품질과 가치평가를 위한 연계 속성 및 평가결과를 항목으로 수용하고 ③ 연계를 위한 속성 및 표준 (DCAT)에 대한 호환성을 고려하였다.

#### 3.2. 데이터 카탈로그

데이터 카탈로그는 여러 관점의 메타데이터의 모음이며 그 메타데이터를 관리하기 위한 기능적 요소를 포함한다.

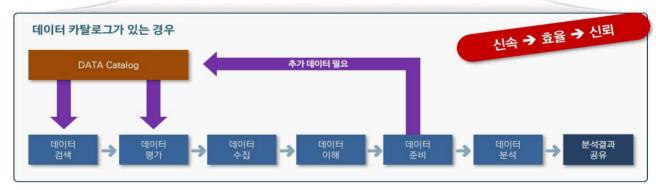
데이터 카탈로그는 메타데이터가 실제 거래소 혹은 유통플랫폼에서 물리적인 관리 서비스의 형태로 구현이 되는 것이다.

데이터 카탈로그는 데이터 이용자가 데이터셋, 데이터셋의 내용, 각 데이터셋의 품질 및 유용성에 대한 가시성을 확보할 수 있게 한다. 이는 데이터를 찾고 이해하는데 소요되는 시간을 단축하고 불필요한 재작업을 방지한다.

데이터 카탈로그 부재의 경우 데이터 수집과 평가를 시행한 후에 데이터에 대한 이해 과정을 거치게 되며 이 과정상에 시행착오와 재작업이 발생한다.

반면 데이터 카탈로그가 존재 시, 카탈로그의 메타데이터를 통해 활용 적합성 판단을 먼저 거친 후 데이터 수집을 진행하기 때문에 부적합한 데이터로 인한 재작업 발생 확률을 현저 히 낮출 수 있다.





[그림-부록-2] 데이터 카탈로그의 필요성

#### 3.3. 데이터 카탈로그와 메타데이터

메타데이터는 데이터 카탈로그의 핵심이며 데이터 카탈로그는 여러 측면의 메타데이터를 포함하고 있다. 데이터 상품 정의를 위한 메타데이터, 검색을 위한 메타데이터, 배포와 유 통을 위한 메타데이터, 계보와 추적을 위한 메타데이터, 연계를 위한 메타데이터, 데이터셋 의 구조에 대한 메타데이터, 데이터셋의 내용과 주제에 대한 메타데이터 등을 포함한다.

메타데이터의 실제 가치는 이해관계자별 요구사항에 대해 메타데이터가 해당 내용을 제공해 주는 것이다.

데이터 구매자는 보고, 분석 및 데이터 과학 작업을 위한 데이터를 찾고 해당 데이터를 평가하여 올바른 데이터셋에서 작동하는지 확인하기 위한 메타데이터가 필요하다.

데이터 공급자는 데이터 수요를 관찰하고, 데이터 구매자의 요구와 관심을 이해하고, 데이터 품질과 가치를 효과적으로 관리하기 위한 메타데이터가 필요하다.

데이터 관리자(거래소 및 플랫폼)는 중요한 데이터를 식별 및 보호하고, 데이터 계보를 관리하고, 데이터에 유통에 대한 신뢰를 확립하기 위한 메타데이터가 필요하다.

#### 3.4. 메타데이터의 필요 정보

메타데이터는 데이터 상품을 정의하면서 다음과 같은 정보를 제공하여야 한다.

- 데이터를 찿고, 이해하고, 평가할 수 있는 정보를 제공하여야 한다.
- 데이터를 분석, 활용하기 위한 데이터의 내용과 주제에 대한 정보를 제공하여야 한다.
- 데이터의 연계 및 융합을 위한 정보를 제공하여야 한다.
- 데이터의 품질 및 가치를 판단하기 위한 정보를 제공하여야 한다.
- 데이터의 유통과 거래의 투명성을 위한 정보를 제공하여야 한다.
- 데이터의 이용과 재생산에 있어 확장성과 제약성에 대한 정보를 제공하여야 한다.





Korea Data Agency

# 메타데이터 체계

#### 4.1. 구성 체계

본 안내서에서 제시하는 이용자 관점의 메타데이터는 그 구성 측면에서 ① 기존 국내 거래 소/플랫폼의 메타데이터 공통항목 수용 ② 품질/가치평가 연계 ③ DCAT 호환성 등을 염두에 두었으며 항목의 선정은 6가지 관점을 통해 구성하였다.

메타데이터는 데이터 상품을 정의하기 위한 비즈니스 메타 성격의 기본정의 항목과 데이터 셋에 대한 기술적(테크니컬) 메타데이터인 데이터명세서(컬럼정의, 코드정의)로 구성한다.

데이터 명세서의 정의 목적은 ① 데이터 셋에 대한 명확한 이해 ② 품질/가치 평가의 기준데이터 제공에 있으며, 타 데이터셋과의 연계 분석을 위해서 필수적으로 제공되어야 하는 정보이다



[그림-부록-3] 이용자 중심 메타데이터의 구성 및 6가지 선정 관점

#### 4.2. 기본정의 항목 구성

메타데이터의 기본정의 항목 구성은 다음과 같다.

기본정의 항목은 총 50개이며 15개의 필수항목으로 구성된다.

기본정의 항목은 ① 이용자를 위한 데이터 상품의 정의 ② 품질/가치 평가의 연계 데이터

③ 카탈로그 간 상호 연계 운용성을 위한 DCAT 호환 속성을 포함함.



[그림-부록-4] 메타데이터-기본정의

#### (1) 기본정의 항목 내역

기본정의 항목의 내역은 다음과 같다.

구분	표준항목	필수여부	가격연계	품질연계	DCAT 호환
	상품카테고리	•	•	•	•
	상품명	•	•		•
	상품설명	•			•
기본속성	과금유형	•			
	상품가격	•			
	가격정책	•			
	사용언어	•			•
	공급자명	•	•		•
	공급자업종				
	공급자홈페이지				
<del>공</del> 급자 속성	공급담당자직급				
	공급담당자명				•
	공급담당자전화번호				•
	공급담당자이메일				•
	상품형태대분류	•		•	•
	상품형태소분류	•		•	•
	확장자		•	•	•
	압축형식		•		•
배포속성	패키지형식		•		•
	제공방식	•	•		
	사이즈		•	•	•
	건수		•	•	
	제공URL			•	•

	활용사례				
	데이터기준		•		•
	키워드	•	•		•
	상 <del>품등록</del> 일자		•		•
	최종수정일자		•	•	•
	상품제공주기	•	•		•
	차기등록예정일자		•		
	데이터시작일자			•	•
부가특성	데이터종료일자			•	•
	데이터주기구분			•	•
	판매시작일자				
	판매종료일자				
	비식별화구분			•	
	부가서비스제공여부	•			
	가공상품여부				
	원천상품제공처				
	원천상품명				
	제공제약사항				
게야기하	라이선스	•	•		•
제약사항	다운로드가능기간				
	데이터사용기간				
	이용자평점				
이오트게	조회건수				
이용통계	판매건수				
	좋아요건수				
평가속성	품질평가지수		•	•	
0/170	가치평가등급		•		

#### (2) 기본정의 항목 작성 기준

기본정의 항목의 세부 작성 기준은 다음과 같다.

속성명	상품카테고리	영문명		prd_catg	
속성구분	기본속성	필수속성	•	타입	텍스트
설명	상품이 속한 카테고리의 명칭				
예시(코드)					
DCAT	dcat:theme	품질연계	•	가격연계	•
비고	O 별도의 관리가 없는 한 상품 상품을 복수의 카테고리에 ·			관계를 가질 수	수 있다. 동일한
속성명	상품명	영문명		prd_nm	
속성구분	기본속성	필수속성	•	타입	텍스트
설명	판매상품의 명 또는 제목				
예시(코드)					
DCAT	dct:title	품질연계		가격연계	
비고	<ul> <li>판매상품의 명은 이용자가 작성하여야 한다.</li> <li>예를들어 "생활 환경 데이터 더 직관적이라 할 수 있음.</li> <li>기본적으로 지역축제에 대한 정보를 제공하고 있음</li> </ul>	l" 보다는 "상t	반기 지역축제 시	시간대 유동인국	P"라는 제목이

속성명	상품설명	영문명		prd_desc			
속성구분	기본속성	필수속성	•	타입	텍스트		
설명	판매상품에 대한 상세 설명						
예시(코드)							
DCAT	dct:description	품질연계		가격연계			
비고	O 상품의 설명은 카탈로그 항목에 명시되지 않는 상품의 기본적인 개요와 특징들을 자						

속성명	과금유형	영문명		chrg_typ		
속성구분	기본속성	필수속성	•	타입	텍스트	
설명	유료, 무료, 협의 등 가격의 유	.형				
예시(코드)	유료 / 무료 / 협의	유료 / 무료 / 협의				
DCAT		품질연계		가격연계	•	
비고	O 무료: 거래소 회원이 이용 다운로드 받을 수 있는 상품 O 유료: 거래소 회원이 결제 받을 수 있는 상품 O 협의: 거래소에 파일리소스 청 접수 후 데이터 상품 제	등 완료 후 파일 느의 등록없이	리소스를 다운 <sup>로</sup> 상품 정보만 입	로드 력되어 있어 <sup>호</sup>		



속성명	상품가격	영문명		prd_prc	
속성구분	기본속성	필수속성	•	타입	숫자
설명	상품의 판매 가격				
예시(코드)	N/A				
DCAT		품질연계		가격연계	
비고	O 무료, 협의 상품일 경우는 [	Default = 0			
속성명	가격정책	영문명		prc plcy	
속성명  속성구분	<b>가격정책</b> 기본속성	영문명 필수속성	•	prc_plcy <b>타입</b>	텍스트
		필수속성	•		텍스트
속성구분	기본속성	<b>필수속성</b> 력정책	●  개 시군구 (VA	타입	텍스트
속성구분 설명	기본속성 비용부과 기준 및 단위 등 가격	<b>필수속성</b> 력정책	● 기개 시군구 (VA	타입	텍스트

속성명	사용언어	영문명		data_lang		
속성구분	기본속성	필수속성	•	타입	텍스트	
설명	데이터셋에 사용된 자연어의 종류 데이터셋의 텍스트(문자)형 메타데이터 항목(예, 상품명, 상품설명 등)에 사용되는 언어 또는 배포되는 데이터셋의 언어를 나타 냅니다.					
예시(코드)						
DCAT	dct:language	품질연계		가격연계		
비고	O Library of Congress(미 의회도서관)정의 속성					
A UPI	2726	OLD EL		ı		

속성명	공급자명	영문명	prvdr_nm		
속성구분	공급자속성	필수속성	•	타입	텍스트
설명	상품을 등록(제공)하는 공급자,	공급기업 및	기관 명		
예시(코드)					
DCAT	dct:publisher	품질연계		가격연계	•
비고					



속성명	공급자업종	영문명		prvdr_biz_typ		
속성구분	공급자속성	필수속성		타입	텍스트	
설명	공급기업 및 기관의 통계청 한	국표준산업분류	로(KSIC)상의 업공	종(대분류)		
예시(코드)	농업, 임업 및 어업(01~03) 광업(05~08) 제조업(10~34) 전기, 가스, 증기 및 공기 조절 공급업(35) 수도, 하수 및 폐기물 처리, 원료 재생업(36~39) 건설업(41~42) 도매 및 소매업(45~47) 운수 및 창고업(49~52) 숙박 및 음식점업(55~56) 정보통신업(58~63)					
DCAT		품질연계		가격연계		
비고	https://kssc.kostat.go.kr:844.	3/ (10차 개정	J)			

속성명	공급자홈페이지	영문명	prvdr_hmpg	
속성구분	공급자속성	필수속성	타입	텍스트
설명	공급기업의 홈페이지 주소			
예시(코드)	http://www.petalake.com/			
DCAT		품질연계	가격연계	
비고				

속성명	공급담당자직급	영문명	р	rvdr_mngr_gı	rd
속성구분	공급자속성	필수속성		타입	텍스트
설명	공급기업의 공급 담당자 직급				
예시(코드)	사원, 주임, 대리, 과장, 차장,	부장, 이사,	티		
DCAT		품질연계		가격연계	
비고					



속성명	공급담당자명	영문명	pr	vdr_mngr_ni	m	
속성구분	공급자속성	필수속성		타입	텍스트	
설명	공급기업의 공급 담당자 명					
예시(코드)	홍길동					
DCAT	dcat:contactPoint	품질연계		가격연계		
비고	O 담당자의 직무별 복수 등록 가능 O 상품을 실제 등록/관리 하는 기술 담당자 O 상품의 판매 관련 담당자					
속성명	공급담당자전화번호	영문명		prvdr_tel		
속성명 속성구분	<b>공급담당자전화번호</b> 공급자속성	영문명 필수속성		prvdr_tel 타입	텍스트	
		필수속성			텍스트	
속성구분	공급자속성	필수속성			텍스트	
속성구분 설명	공급자속성 공급기업의 공급 담당자 전화번	필수속성			텍스트	

속성명	공급담당자이메일	영문명		prvdr_email		
속성구분	공급자속성	필수속성		타입	텍스트	
설명	공급기업의 공급 담당자 이메일 주소					
예시(코드)	gildong.hong@petalake.com					
DCAT	dcat:contactPoint	품질연계		가격연계		
비고	O 담당자의 직무별 복수 등록 O 상품을 실제 등록/관리 하는 O 상품의 판매 관련 담당자					
속성명	상품형태대분류	영문명		prdtyp_lclsfy		
속성구분	배포속성	필수속성	•	타입	텍스트	
설명	상품형태 대분류 (정형/반정형/	'비정형)				
예시(코드)	정형 / 반정형 / 비정형					
DCAT	dcat:mediaType	품질연계	•	가격연계		

#### 이 확장자 기준 자동판단 가능 비고

O 압축파일 및 식별 불가 시 공급자 입력



속성명	상품형태소분류	영문명		prdtyp_sclsfy		
속성구분	배포속성	필수속성	•	타입	텍스트	
설명	상품형태 소분류 (문서/이미지/음성/영상/래스터/벡터/공간속성/)					
예시(코드)	문서 / 이미지 / 음성 / 영상 /	'래스터 / 벡터	터 / 공간속성			
DCAT	dcat:mediaType	품질연계		가격연계		
비고	O 확장자 기준 자동판단 가능 O 압축파일 및 식별 불가 시 공급자 입력					
	확장자 영문명 ext nm					
속성명	확장자	영문명		ext_nm		
속성명 속성구분	<b>확장자</b> 배포속성	영문명 필수속성	•	ext_nm 타입	텍스트	
			•		텍스트	
속성구분	배포속성	필수속성	• ML		텍스트	
속성구 <del>분</del> 설명	배포속성 제공되는 데이터파일의 확장자	필수속성	• ML		텍스트 <b>●</b>	

속성명	압축형식	영문명		cmprs_frmt	
속성구분	배포속성	필수속성	•	타입	텍스트
설명	데이터셋 파일의 압축형식				
예시(코드)	zip				
DCAT	dcat:compressFormat	품질연계		가격연계	•
비고	O IANA에서 정의한 유형 https://www.iana.org/assign	ments/media	-types/media-t	ypes.xhtml	

속성명	패키지형식	영문명		pkg_frmt	
속성구분	배포속성	필수속성	•	타입	텍스트
설명	데이터셋 파일의 패키지형식				
예시(코드)	tar				
DCAT	dcat:packageFormat	품질연계		가격연계	•
비고	O IANA에서 정의한 유형 https://www.iana.org/assign	ments/media	-types/media-t	ypes.xhtml	



속성명	제공방식	영문명		distrib_typ		
속성구분	배포속성	필수속성	•	타입	텍스트	
설명	이용자에게 상품이 제공되는 물리적인 인터페이스 방식					
예시(코드)	API, Download, eMail, 등					
DCAT		품질연계		가격연계	•	
비고						
	사이즈 영문명 size					
속성명	사이즈	영문명		size		
속성명 속성구분	<b>사이즈</b> 배포속성	영문명 필수속성	•	size <b>타입</b>	텍스트	
		필수속성	•		텍스트	
속성구분	배포속성	필수속성	•		텍스트	
속성구분 설명	배포속성	필수속성	•		텍스트 <b>●</b>	

속성명	건수	영문명		row_cnt	
속성구분	배포속성	필수속성	•	타입	텍스트
설명	제공되는 데이터셋 파일내의 더	베이터 행 수			
예시(코드)					
DCAT		품질연계	•	가격연계	•

헤더가 포함된 경우 헤더를 제외한 실제 데이터 건수

비고

속성명	제공URL	영문명		distrib_url	
속성구분	배포속성	필수속성	•	타입	텍스트
설명	상품의 다운로드나 API 연동 등	등 판매자가 수	요자에게 상품을	을 제공하기 위	한 URL
예시(코드)					
DCAT	dcat:downloadURL	품질연계	•	가격연계	
비고					



속성명	활용사례	영문명		use_case			
속성구분	부가특성	필수속성	•	타입	텍스트		
설명	데이터 상품의 활용사례						
예시(코드)	[로그인 시스템] 결제 직전 악성 IP주소 정보 데이터를 분석하여 부정적인 결과가 나오면 사용자에게 별도의 notification을 주거나 결제 보류를 시킬 수 있습니다.						
DCAT		품질연계		가격연계			
비고	데이터 상품의 활용사례를 제공	당하여 이용자의	리 연계 활용 가	치를 제공			
	데이터기준 영문명 data_criteria						
속성명	데이터기준	영문명		data_criteria			
속성명 속성구분	<b>데이터기준</b> 부가특성	영문명 필수속성	•	data_criteria 타입	텍스트		
		필수속성	•				
속성구분	부가특성	필수속성	•				
속성구분 설명	부가특성 데이터 상품의 분석 차원 기준	필수속성	•				

속성명	키워드	영문명		keyword		
속성구분	부가특성	필수속성	•	타입	텍스트	
설명	상품 검색을 위한 키워드, 해시태그					
예시(코드)	상권분석, 유동인구					
DCAT	dcat:keyword	품질연계		가격연계	•	
비고	O 키워드는 이용자가 데이터 로 최대한 데이터 상품의 나					
속성명	상 <del>품등</del> 록일자	영문명		issue_dt		
속성구분	부가특성	필수속성	•	타입	텍스트	
설명	거래소에 상품을 등록한 일자					
예시(코드)						
DCAT	dct:issued	품질연계		가격연계	•	

비고



속성명	최종수정일자	영문명		modify_dt		
속성구분	부가특성	필수속성	•	타입	텍스트	
설명	상품정보 최종 수정일자					
예시(코드)						
DCAT	dct:modified	품질연계	•	가격연계	•	
비고	O 상품 정보에 대한 최종수정일자로 공급자가 등록하는 항목이 아닌 시스템 Timestamp를 통해 생성 O 최종수정일자는 데이터셋의 내용에 대한 수정일과 반드시 일치 하지는 않는다.					
속성명	상품제공주기	영문명		issue_period		
속성명 속성구분	<b>상품제공주기</b> 부가특성	영문명 필수속성	•	issue_period 타입	텍스트	
			•		텍스트	
속성구분	부가특성	필수속성	•		텍스트	
속성구분 설명	부가특성 상품의 제공(등록) 주기	필수속성	•		텍스트 <b>●</b>	

속성명	차기등록예정일자	영문명		nxt_issue_dt	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	상품의 제공주기에 따른 차기등	등록예정일			
예시(코드)	'20201130'				
DCAT		품질연계		가격연계	•
			1		

비고 상품의 등록일자에 상품제공주기 만큼의 기간을 가산하여 차기증록예정일을 산정함.

속성명	데이터시작일자	영문명		ds_st_dt	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	데이터셋 내 데이터의 시작일지	ŀ			
예시(코드)	3년간(2017 ~ 2019)의 시계( '20170101'	결 데이터를 월	실별로 제공하는	데이터셋일 경	우
DCAT	dct:temporal	품질연계	•	가격연계	
비고					



속성명	데이터종료일자	영문명		ds_ed_dt	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	데이터셋 내 데이터의 종료일지	4			
예시(코드)	3년간(2017 ~ 2019)의 시계( '20191231'	열 데이터를 월	l별로 제공하는	데이터셋일 경	우
DCAT	dct:temporal	품질연계	•	가격연계	
비고					
속성명	데이터주기구분	영문명	C	data_period_c	d
솔성구 <del>분</del>	부가 <del>특</del> 성	필수속성		타인	텐스트

속성명	데이터주기구분	영문명	C	data_period_c	d
속성구분	부가특성	필수속성	•	타입	텍스트
설명	시계열 데이터일 경우 데이터의	기 주기(간격)			
예시(코드)	3년간(2017 ~ 2019)의 시계(	열 데이터를 월	실별로 제공하는	데이터셋일 경	우 주기는 '월'
DCAT	dcat:temporalResolution	품질연계	•	가격연계	
비고					

속성명	판매시작일자	영문명		sales_st_dt	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	해당 상품의 판매 시작일자				
예시(코드)	'20201101'				
DCAT		품질연계		가격연계	
비고					

속성명	판매종료일자	영문명		sales_ed_dt	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	해당 상품의 판매 종료일자				
예시(코드)	'20201131'				
DCAT		품질연계		가격연계	
비고					



속성명	비식별화구분	영문명		non_ident_cd	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	가명, 익명 처리 등 개인정보어	대한 비식별	화 처리 구분		
예시(코드)	가명처리 / 익명처리 / 해당없음	0			
DCAT		품질연계	•	가격연계	
비고					
속성명	부가서비스제공여부	영문명		add_svc_yn	
속성명 속성구분	<b>부가서비스제공여부</b> 부가특성	영문명 필수속성	•	add_svc_yn 타입	텍스트
		필수속성	● 기가능한 부가서!	타입	
속성구분	부가특성	필수속성	●  가능한 부가서!	타입	
속성구분 설명	부가특성 판매상품의 후가공, 제공방식의	필수속성	●	타입	

속성명	가공상품여부	영문명		prcs_prd_yn	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	판매상품이 타사의 원천 상품으	으로 부터 가공	에 의해 만들어	진 상품 여부	
예시(코드)	Y / N				
DCAT		품질연계		가격연계	

단하는 기준이 될 수 있음.

O 또한 데이터 유통의 투명성이나 추적성을 위해 명시

비고

O 가공 후 (재)판매 상품일 경우는 원천 상품에 따라 가공상품의 품질 및 신뢰도를 판

속성명	원천상품제공처	영문명		org_prd_prvd	r
속성구분	부가특성	필수속성	•	타입	텍스트
설명	가공상품의 경우 원천상품의 제공처				
예시(코드)	상품명 : 이동통신 사용자로 추 원천상품제공처 : KT	녹계된 인구통계	정보		
DCAT		품질연계		가격연계	
비고	<ul><li>가공 후 (재)판매 상품일 경단하는 기준이 될 수 있음.</li><li>또한 데이터 유통의 투명성</li></ul>			상품의 품질 및	및 신뢰도를 판



속성명	원천상품명	영문명		org_prd_nm	
속성구분	부가특성	필수속성	•	타입	텍스트
설명	가공상품의 경우 원천상품의 상품명				
예시(코드)	상품명 : 이동통신사용자로 추기원천상품명 : 50cell LTE 유동연		정보		
DCAT		품질연계		가격연계	
비고	O 가공 후 (재)판매 상품일 경단하는 기준이 될 수 있음. O 또한 데이터 유통의 투명성이			상품의 품질 및	및 신뢰도를 판
속성명	제공제약사항	영문명		prvdr_limit	
속성명 속성구분	<b>제공제약사항</b> 제약사항	영문명 필수속성	•	prvdr_limit 타입	텍스트
		필수속성	•		텍스트
속성구분	제약사항	<b>필수속성</b> 항을 기술 사용 가능 용 특별 공급 나 재판매, 제3	자 공유, 배포 ¦	<b>타입</b> 불가능	
속성구분 설명	제약사항 상품 판매(제공) 측면의 제약사 - 데이터 바우처 사업 목적 내 - 중소기업, 1인기업, 소상공인 - 사업목적 외 타 목적 사용 및 - 사업 목적 내 가공 재판매, 7	<b>필수속성</b> 항을 기술 사용 가능 용 특별 공급 나 재판매, 제3	자 공유, 배포 ¦	<b>타입</b> 불가능	

속성명	라이선스		영문명		license	
속성구분	제약사항		필수속성	•	타입	텍스트
설명	라이선스 정책에 대한	사항을	기술			
예시(코드)	CCO CC BY CC BY-SA CC BY-NC CC BY-NC-SA CC BY-ND CC BY-NC-ND	저작자 저작자 저작자 저작자 저작자	를게 이용가능 나 표시 나 표시-동일조 나 표시-비영리 나 표시-비영리 나 표시-변경금 나 표시-비영리	-동일조건변경허 지	락	
DCAT	dct:license		품질연계		가격연계	•
비고	O CCL(Creative Com 이용할 수 있도록 ā			작권자가 지정힌	· 조건에 한하(	여 그 저작물을
속성명	다운로드가능기간	<u>'</u>	영문명	d	lownload_terr	n
속성구분	제약사항		필수속성	•	타입	숫자
설명	상품의 구매 후 다운로	드 가능	기간			
예시(코드)	7 : 구매일자로 부터 7	7일간 다	운로드 가능			
DCAT			품질연계		가격연계	
비고	O 상품 구매일자 기준	: 다운로	드 가능 기간을	을 일수로 표시		



속성명	데이터사용기간	영문명	C	data_use_term	1
속성구분	제약사항	필수속성	•	타입	숫자
설명	상품의 구매 혹은 다운로드 후	사용가능 기간	<u>'</u>		
예시(코드)	30 : 다운로드 후 30일간 사용	강가능			
DCAT		품질연계		가격연계	
비고	O 상품 다운로드 일자 기준 시	h용 기간을 일	수로 표시		
속성명	이용자평점	영문명	L	usr_givn_grade	2
속성명 속성구분	<b>이용자평점</b> 이용통계	영문명 필수속성	•	usr_givn_grade 타입	e 숫자
			•		
속성구분	이용통계		•		
속성구분 설명	이용통계 실 이용자 평점		•		



이 해당 상품의 조회 건수를 합산하며 동일한 이용자에 의한 반복 조회는 1건으로 처리

속성명	판매건수	영문명		sales_cnt	
속성구분	이용통계	필수속성	•	타입	숫자
설명	상품이 판매(다운로드)된 건수				
예시(코드)	15 (건)				
DCAT		품질연계		가격연계	
비고	O 해당 상품의 판매(다운로드)	건수를 합산			



속성명	좋아요건수	영문명		like_cnt	
속성구분	이용통계	필수속성	•	타입	숫자
설명	좋아요 등록 건수				
예시(코드)	135 (건)				
DCAT		품질연계		가격연계	
비고	O 해당 상품에 대한 이용자의	좋아요 등록	건수		
속성명	품질평가지수	영문명		q_eval_ind	
속성명 속성구분	<b>품질평가지수</b> 평가속성	영문명 필수속성	•	q_eval_ind 타입	숫자
		필수속성	•		숫자
속성구분	평가속성	필수속성	•		숫자
속성구분 설명	평가속성 품질평가 모형에서 산출된 상품	필수속성	•		숫자

	ı			
P	2			
S	F	=		
r	9	h	2	1
P	١	Č	ų	ı
ι	)	Ľ	Z	
ъ.	d	w		۰

메타데이터 체계

속성명	가치평가등급	영문명		v_eval_grd	
속성구분	평가속성	필수속성	•	타입	텍스트
설명	가치평가모형에서 산출된 상품의	의 평가등급			
예시(코드)	최상 / 상 / 중 / 중하 / 하				
DCAT		품질연계		가격연계	•
비고	O 가치평가 모형에서 산출된 ·	등급으로 5단2	예 중 하나의 값	을 가진다.	



#### 4.3. 데이터 명세서 구성

데이터 명세서는 컬럼정의서와 코드정의서로 구성되며 세부 항목은 다음과 같음. 컬럼정의서는 데이터셋의 개별항목(컬럼)에 대한 기술(테크니컬) 메타데이터 이며 컬럼의 형식과 내용에 대한 설명을 제공 함.

코드정의서는 컬럼정의서 상에 명시된 항목 중 코드성 항목에 대한 세부 코드 내역을 정의 하며 데이터셋의 해당항목 데이터는 반드시 정의된 코드 값 내에서 발생하여야 함. 데이터셋의 품질평가를 하기 위한 필수 메타데이터 임.



[그림-부록-5] 메타데이터-데이터명세서

### (1) 컬럼정의서 작성 기준

컬럼정의서의 개별 항목별 작성 기준은 다음과 같음.

속성명	컬럼명	영문명		col_nm	
속성구분	컬럼정의	필수속성	•	타입	텍스트
설명	컬럼에 대한 명칭				
예시					
품질연계	•	가격연계		•	
비고	O 컬럼명은 통상 물리적인 Da 제공하는 데이터셋의 항목에 O 다만, 통상의 표준화 명명 불이는 것이 직관적인 컬럼 O 예를들어 코드성 컬럼의 경한 컬럼의 경우 "_RT" 등으	대한 명칭 C 구칙에서 정의 이 유형을 판단 우 "_CD"로 등	기므로 그 명명어하는 방식대로 <sup>†</sup> 단함에 있어 바림 금액 컬럼의 경우	에 특별한 제약/ 컬럼 도메인 유 남직함. P "_AMT, _B/	사항은 없음. 2형별로 어미를 AL", 비율에 대



속성명	컬럼한글명	영문명		col_ko_nm		
속성구분	컬럼정의	필수속성	•	타입	텍스트	
설명	컬럼에 대한 한글 명칭					
예시						
품질연계	•	가격연계		•		
비고	<ul> <li>O 컬럼한글명은 컬럼명에 대응하는 한글 명으로 통상의 표준화 명명규칙에서 정의하는 방식대로 컬럼 도메인 유형별로 어미를 붙이는 것이 직관적인 컬럼의 유형을 판단함에 있어 바람직함.</li> <li>O 예를들어 코드성 컬럼의 경우 "코드, 구분코드"로 금액 컬럼의 경우 "금액, 잔액", 비율에 대한 컬럼의 경우 "율, 률" 등으로 끝나도록 명명하는것이 좋은 명명 방식임.</li> </ul>					
속성명	데이터타입	영문명		col_type		
속성구분	컬럼정의	필수속성	•	타입	텍스트	
설명	컬럼 데이터의 자료형식					
예시	char, varchar, number, date	등				
품질연계	•	가격연계				
비고	O 데이터 타입은 기본적으로 /	ANSI 기준을 <sup>I</sup>	따르도록 함.			

속성명	길이	영문명	col_len	
속성구분	컬럼정의	필수속성	타입	텍스트
설명	컬럼 데이터의 길이			
예시	100, 18,3			
품질연계	•	가격연계		
비고	O 컬럼의 길이는 데이터타입어 O 데이터 타입에 따라 길이를			

속성명	PK여부	영문명		pk_yn	
속성구분	컬럼정의	필수속성	•	타입	텍스트
설명	컬럼이 PK(Primary Key) 속성	인지 여부			
예시	Y / N				
품질연계	•	가격연계		•	
비고					



속성명	NotNull여부	영문명		nn_yn	
속성구분	컬럼정의	필수속성	•	타입	텍스트
설명	컬럼이 NotNull(Null 값을 허용	용하지 않는) Q	인지 여부		
예시	Y / N				
품질연계	•	가격연계		•	
비고	비고         O 'Y' 일 경우는 NotNull, 즉 Null을 허용하지 않음         O PK 항목일 경우는 반드시 'Y'(NotNull) 이어야 함				
속성명	컬럼설명	영문명		col_desc	
속성구분	컬럼정의	필수속성	•	타입	텍스트
설명	컬럼의 내용에 대한 상세 설명				
예시	1) 인당매출금액 : 1인당 매출 값 (총매출금액 / 총구매자수) 2) 건당매출금액 : 구매 1건당 나눈 값 (총매출금액 / 총구매장	방 매 <del>출</del> 금액 이			
품질연계	•	가격연계		•	
비고	O 컬럼설명은 컬럼의 내용을 여 컬럼이 집계나 연산에 의해 O 데이셋 내의 다른 컬럼으로 리 연산도 기술 함.	생성되는 경우	우 상세한 연산식	시이나 집계내용	을 기술 함.

속성명	코드분류명	영문명		cd_grp_nm	
속성구분	컬럼정의	필수속성	•	타입	텍스트
설명	코드성 컬럼일 경우 해당 코드	의 코드분류명			
예시	성별구분코드, 연령대코드, 업종	<u> </u>			
품질연계	•	가격연계		•	
비고	O 코드성 컬럼일 경우 코드분· O 코드분류명은 코드정의서의				

속성명	도메인그룹	영문명	domain_grp_nm			
속성구분	컬럼정의	필수속성	•	타입	텍스트	
설명	컬럼이 속한 도메인그룹명					
예시	금액 / 내용 / 명 / 번호 / 수 / 시각 / 율 / 일자 / 코드					
품질연계	•	가격연계		•		
비고						



속성명	도메인	영문명	domain_nm				
속성구분	컬럼정의	필수속성	•	타입	텍스트		
설명	컬럼이 속한 도메인명						
예시	〈 금액 도메인 그룹일 경우 〉 금액 / 료 / 비용 / 세액 / 수수료 / 외화금액 / 이자 / 잔액 / 적수 / 평잔 등						
품질연계	•	가격연계	•				
비고	O 도메인명으로 컬럼명이 끝나도록 일치시켜서 명명						
	유효기준 영문명 valid_ref						
속성명	유효기준	영문명		valid_ref			
속성명 속성구분	<b>유효기준</b> 컬럼정의	영문명 필수속성		valid_ref 타입	텍스트		
					텍스트		
속성구분	컬럼정의	필수속성			텍스트		
속성구분 설명	컬럼정의 지정된 도메인 유효값의 기준	필수속성	)		텍스트		

속성명	연계키		영문명	link_key			
속성구분	컬럼정의		필수속성		타입	텍스트	
설명	데이터셋간의 상호 연계를 위한 연계키						
	영역		표준연계키				
	사람	성, 연령, 연령대, 이메일, 거주지 주소, 우편번호					
예시	기업	사업체 명칭, 사업체 전화번호, 사업체 주소, 법인번호, 업종					
	시간	간 시간, 시각, 날짜, 연도					
	공간	주소, 좌표, 행정구	주소, 좌표, 행정구역, 국가기초구역, 건물(집합건물, 빌딩)				
	(출처) 통계청발간 통계의 창 2019 겨울호						
품질연계	● 가격연계 ●						
비고	O 정보화 진흥원에서 정의한 표준 연계키를 각 컬럼에 대해 지정. O 표준 연계키를 지정하는 컬럼의 내용은 반드시 코드 등, 정형화된 형식을 갖춘 데이터 이어야 함.						



#### (2) 코드 정의서 작성 기준

코드정의서의 개별 항목별 작성 기준은 다음과 같음.

속성명	코드분류명	영문명	cd_grp_nm				
속성구분	코드정의	필수속성	•	타입	텍스트		
설명	코드의 코드분류명						
예시	성별구분코드, 연령대코드, 업종코드						
품질연계	•	가격연계					
비고	O 코드분류명은 컬럼정의서의	코드분류명과	일치 함.				
속성명	코드값	영문명	cd_val				
속성구분	코드정의	필수속성	•	타입	텍스트		
설명	코드의 실제 발생 값(인스턴스)						
예시	코드분류명 : 성별구분코드 코드값 : 'M', 'F'						
품질연계	•	가격연계		•			
비고	O 코드 값은 해당 코드그룹을 하는 것임.	사용하는 데여	기터셋 컬럼의 별	날생 가능한 값의	의 범위를 정의		

속성명	코드명	영문명		cd_nm	
속성구분	코드정의	필수속성	•	타입	텍스트
설명	코드 값에 대한 명칭				
예시	코드분류명 : 성별구분코드 코드값 : 'M', 'F' 코드명 : '남성', '여성'				
품질연계	•	가격연계			
비고	O 코드명은 데이터셋의 코드성 사용 함.	성 컬럼의 발생	! 코드값을 UI등	·에 그 의미를	표현하기 위해

속성명	코드설명	영문명		cd_desc		
속성구분	코드정의	필수속성	•	타입	텍스트	
설명	코드 값에 대한 상세 설명					
예시	코드분류명 : 연령대코드 코드값 : '10', '20' 코드명 : '10대', '20대' [ 코드설명 ] 10대 : 연령 10 ~ 19세 20대 : 연령 20 ~ 29세 					
품질연계	•	가격연계				
비고	O 코드 설명은 코드명 만으로는 코드값의 의미가 명확치 않은 경우, 예를들어 구간 값을 가지는 코드에 대해 상세한 설명을 기술 함.					

### 2020 종합안내서 데이터 거래 지원 가이드라인

발행일 2020년 12월

**발행인** 민기영

발행처 한국데이터산업진흥원

(04513) 서울시 중구 세종대로9길 42 부영빌딩 8층 TEL: 02-3708-5300 / FAX: 02-318-5040 www.kdata.or.kr / www.datastore.or.kr

본 안내서와 관련한 문의는 한국데이터산업진흥원 유통기반실 유통기획팀으로 연락해 주시기 바랍니다.







