



# 데이터 요소의 표준화(V)

장윤진

한국체육과학연구원

### 차례

- I. 총론
- II 데이터 요소의 생성과 표준화를 위한 프레임워크
- III. 도메인 식별을 위한 개념의 분류
- IV. 기본 데이터 요소 속성
- ▶ V. 데이터 정의의 형식화를 위한 규칙과 지침
- VI. 데이터 요소를 위한 명명 및 식별 원칙
- VII. 데이터 요소의 등록

## V. 데이터 정의의 형식화를 위한 규칙과 지침

### 1. 적용범위

데이터 요소 정의를 구축하기 위한 규칙과 지침에서는 데이터 요소의 의미적 구조만 제시된다. 즉, 정의들을 포맷팅하기 위한 명세서는 이 표준의 목적에 비추어 불필요한 것으로 판단된다.

이 표준에서 규칙과 지침은 비록 데이터 요소에 대한 것이지만, 이들은 또한 entity types, entities, 관계, 속성, 객체형(클래스), 객체, 세그먼트, 구성요소, 코드 엔트리, 메세지 등과 같은 데이터 구조물의 다른 타임을 위한 정의를 형식화하는 데 적용될 수도 있다.

이 표준에서 정의적 규칙과 지침은 용어 사전과 언어 사전에서 발견되는 용어상의 정의에 항상 적용되는 것은 아니다. 언어 사전에서 적용하는 규칙과 데이터 사전에서 적용하는 규칙에는

차이가 있다. 예를 들어, 언어 사전에서는 용어가 복합적으로 여러 의미를 가지도록 정의될 수 있지만, 데이터 사전에서는 유일하게 하나의 의미를 가지도록 정의되어야 한다.

많은 데이터 요소 정의는 자체로 또다시 정의가 필요한 용어들을 포함하고 있다. 용어들중 어떤 것은 다른 산업 분야에서는 다르게 정의될 수 있다. 그러므로, 대부분의 데이터 사전에서는 각 정의들에서 사용된 관련 용어들로 용어 사전을 만들어 둘 필요가 있다. 각 용어의 용법 영역은 그 용어 사전에서 확인하면 된다.

### 2. 정의

이 표준의 목적에 따라, 다음과 같이 용어들을 정의한다.

- 속성 : 객체나 실체(Entity)의 특성.
- 개념 : 객체의 집합에 공통된 특성을 근거로 추출하여 구성된 생각의 단위.

- **데이터** : 인간이나 자동적인 방법에 의한 통신, 해석, 또는 처리에 적합하도록 형식화된 방식으로 나타낸 사실, 개념, 또는 명령들의 표현.
- **데이터 사전** : 다른 데이터의 사용과 구조에 대해 언급한 데이터를 위해 사용된 데이터 베이스 ; 즉, 메타 데이터를 저장한 데이터베이스. 데이터 요소 사전 참조
- **데이터 요소** : 정의, 식별, 표현 및 허용가능한 값들을 속성의 집합을 사용하여 규정한 데이터 단위.
- **데이터 요소 사전** : 관련된 모든 데이터 요소를 열거하고 정의한 정보 자원. 데이터 사전과 유사함.
- **정의** : 사람이나 사물 및 관련 계층의 핵심적 성격을 표현하는 단어나 구.
- **도메인** : 하나의 속성에서 가능한 데이터 값의 집합.
- **이름** : 인간을 위해 객체와 개념을 식별하는 주요 수단.

른 개념에서 정의된 개념과 차이가 있어야 한다.

2) 단수(The Singular)로 진술되어야 한다. 데이터 정의에서 표현된 개념은, 개념 자체가 복수인 경우 이외에는, 단수로 표현되어야 한다.

3) 그 개념이 아닌 것뿐만 아니라, 그 개념이 무엇인지(What the concept is, not only what it is not)를 진술해야 한다. 데이터 의미에서 무엇이 포함되는지를 규정하지 않는 정의는 적절한 것이 아니다.

4) 서술적인 구나 문장으로 진술되어야 한다. 명확히 정의하기 위해 구나, 필요하다면, 완전한 문장을 사용할 필요가 있다. 단순히 한 두개의 동의어로 나타내는 것으로는 불충분하다.

예) "Agent Name"

좋은 정의 : Name of party authorized to act on behalf of another party

미흡한 정의 : Representative

5) 보편적으로 이해되는 약어만 포함해야 한다. 보통 약어나 두문자의 의미는 사용되는 환경에 따라 해석되므로, 보편적으로 사용되는 약어 외에는 오해의 여지가 없도록 풀어써야 한다.

6) 다른 데이터 요소의 정의나 이에 근거한 또다른 개념이 포함되지 않게 표현되어야 한다.

부차적인 데이터 요소나 이에 관련된 개념의 정의는 기본적인 데이터 요소의 정의에 나오지 않아야 한다. 용어의 정의는 관련된 용어사전에서 제공되어야 한다. 만일 추가적인 정의가 필요하다면 기본적 정의의 본문 끝에 주석으로 붙이거나, 용어 사전에서 별도의 항목으로 취급되어야 한다. 관련된 정의는 관계 속성(예를 들면 상호 참조)을 통해 접근되어야 한다.

### 3. 데이터 정의 규칙과 지침

각 규칙과 지침들은 실질적으로 적용할 경우에는 특정 환경과 문맥에 따라 다르게 구현될 수 있으나, 규칙과 지침 그 자체는 직관적으로 수긍할 수 있는 것이다. 이들을 열거하고 간단히 설명하면 다음과 같다.

#### 3.1 규칙

데이터 정의는 :

1) (그것이 있는 어떤 데이터 사전내에서도) 유일(Unique)해야 한다.

각 정의는 그 개념의 특이성이 손실되지 않도록, 사전내의 다른 정의와 구별이 가능해야 한다. 그 정의내에서 표현된 하나 이상의 특징들이 다





# 데이터베이스 리포트①

데이터 요소의 표준화(V)

예) "Sample Type Code"

잘된 정의 : 표본의 종류를 입증하는 코드

잘못된 정의 : 표본의 종류를 입증하는 코드.  
이 정의에서 다시 표본을 정의하는 경우.

5) 합리화, 기능적 용법, 도메인 정보 또는 절차적 정보 없이 표현되어야 한다.

이들 정보가 간혹 필요하다 해도, 해당 정의의 목적에 관계없는 정보가 포함되므로 적절한 정의라고 할 수 없다.

- 주어진 정의에 대한 합리화는 정의 부분에서 제외되어서는 안된다. 예를 들어, 킬로미터 대신에 마일즈를 데이터요소가 사용한다 해도 그 이유를 정의에서 지적해서는 안된다.
- "이 데이터 요소는.. 로 사용해서는 안된다."와 같은 기능적 용법은 정의에 포함되지 않는 것이 적절하다.
- 절차적 측면에서 "이 데이터 요소는 데이터 요소 ××와 결합하여 사용된다."와 같은 말은 정의에서 나타내지 않아야 한다.

예) "Data Field Label"

이 자체에 대한 정의외에 field의 기능적 용법에 대한 주석은 정의에서 제외되어야 한다.

6) 순환논리는 피해야 한다.

다음은 잘못된 정의의 예이다.

- "Employee ID Number" : Employee에게 할당된 번호
- "Employee" : Employee ID number와 상응한 사람

7) 관련된 정의에 대해 동일한 용법과 일관된 논리적 구조를 사용해야 한다.

유사하거나 연관된 정의에서 사용된 용어나 구문은 다음 예와 같이 동일해야 한다.

예) "Goods Dispatch Date" : 상품이 해당 자에 의해 발송된 날짜

"Goods Receipt Date" : 상품이 해당자에 의해 수령된 날짜 **DC**

<다음호에는 데이터요소를 위한 명명 및 식별원칙을 게재하고자 한다.>

## 3.2 지침들

데이터 정의는 :

1) 개념의 핵심적 의미를 나타내야 한다.

표현된 개념의 모든 기본적인 특성이 문맥의 특수성과 관계된 수준에서 정의에 나타나야 한다. 핵심이 아닌 특성은 포함하지 말아야 한다. 어느 수준까지 포함할 것인가는 시스템 사용자와 환경의 요구에 따른다.

2) 분명하고 애매하지 않아야 한다.

정의된 개념의 정확한 의미와 해석은 그 정의로부터 명백히 드러나야 한다. 정의는 오직 하나의 해석만 가능하도록 분명해야 한다.

3) 간결해야 한다.

정의는 간결하고 포괄적이어야 한다. 관계없는 한정구들은 사용하지 말아야 한다.

예) "Character Set Name"

잘된 정의 : 데이터가 해석되는 소리글자 또는 뜻글자 기호의 집합에 주어진 이름.

이 정의외에 관계없는 말들을 추가하여 정의할 필요는 없다.

4) 독자적이어야 한다.

개념의 의미는 정의에서 분명히 나타내어야 한다. 추가 설명이나 참조가 정의의 뜻을 이해하는데 필요하지 않아야 한다.

예) "School Location City Name"

잘된 정의 : 학교가 위치한 도시의 이름

잘못된 정의 : "school site"를 보라