

TCVN 10249-100:2013

ISO/TS 8000-100:2009

Xuất bản lần 1

**CHẤT LƯỢNG DỮ LIỆU –
PHẦN 100: DỮ LIỆU CÁI:
TRAO ĐỔI DỮ LIỆU ĐẶC TRƯNG: TỔNG QUAN**

Data quality –

Part 100: Master data: Exchange of characteristic data: Overview

Mục lục	Trang
Lời nói đầu	4
1 Phạm vi áp dụng.....	5
2 Tài liệu viện dẫn.....	6
3 Thuật ngữ và định nghĩa	6
4 Thuật ngữ viết tắt.....	6
5 Dữ liệu cái	6
6 Kiến trúc dữ liệu cho dữ liệu cái.....	9
7 Mô hình dữ liệu mức cao.....	10
8 Tổng quan về bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái.....	14
Phụ lục A_(quy định)_Nhận diện tài liệu	16
Phụ lục B_(tham khảo)_Danh mục hạng mục	17
Thư mục tài liệu tham khảo	20

Lời nói đầu

TCVN 10249-100:2013 hoàn toàn tương đương với ISO/TS 8000-100:2009.

TCVN 10249-100:2013 do Ban Kỹ thuật tiêu chuẩn quốc gia TCVN/JTC1 Công nghệ Thông tin biên soạn, Tổng cục Tiêu chuẩn Đo lường Chất lượng đề nghị, Bộ Khoa học và Công nghệ công bố.

Bộ tiêu chuẩn TCVN 10249 (ISO 8000) *Chất lượng dữ liệu* gồm các tiêu chuẩn sau:

- TCVN 10249-1:2013 (ISO/TS 8000-1:2011), *Phần 1: Tổng quan.*
- TCVN 10249-2:2013 (ISO 8000-2:2012), *Phần 2: Từ vựng.*
- TCVN 10249-100:2013 (ISO/TS 8000-100:2009), *Phần 100: Dữ liệu cái – Trao đổi dữ liệu đặc trưng – Tổng quan.*
- TCVN 10249-110:2013 (ISO 8000-110:2009), *Phần 110: Dữ liệu cái – Trao đổi dữ liệu đặc trưng – Cú pháp, mã hóa ngữ nghĩa và sự phù hợp với đặc tả dữ liệu.*
- TCVN 10249-120:2013 (ISO/TS 8000-120:2009), *Phần 120: Dữ liệu cái – Trao đổi dữ liệu đặc trưng – Xuất xứ.*
- TCVN 10249-130:2013 (ISO/TS 8000-130:2009), *Phần 130: Dữ liệu cái – Trao đổi dữ liệu đặc trưng – Độ chính xác.*
- TCVN 10249-140:2013 (ISO/TS 8000-140:2009), *Phần 140: Dữ liệu cái – Trao đổi dữ liệu đặc trưng – Tính đầy đủ.*
- TCVN 10249-150:2013 (ISO/TS 8000-150:2011), *Phần 150: Dữ liệu cái – Khung quản lý chất lượng.*
- TCVN 10249-311:2013 (ISO/TS 8000-311:2012), *Phần 311: Hướng dẫn ứng dụng chất lượng dữ liệu sản phẩm về hình dáng (PDQ-S).*

Chất lượng dữ liệu –

Phần 100: Dữ liệu cái: Trao đổi dữ liệu đặc trưng: Tổng quan

Data quality –

Part 100: Master data: Exchange of characteristic data: Overview

1 Phạm vi áp dụng

Tiêu chuẩn này nêu tổng quan về chất lượng dữ liệu cái cho chuỗi tiêu chuẩn trong bộ tiêu chuẩn TCVN 10249 (ISO 8000).

Phạm vi của chuỗi tiêu chuẩn về chất lượng dữ liệu cái trong bộ tiêu chuẩn TCVN 10249 (ISO 8000) bao gồm:

- các khía cạnh quy định dữ liệu cái của hệ thống quản lý chất lượng;
- các thang đo chất lượng dữ liệu cái.

Phương pháp tiếp cận chuỗi tiêu chuẩn về chất lượng dữ liệu cái trong bộ tiêu chuẩn TCVN 10249 (ISO 8000) xác định chất lượng dữ liệu:

- từ “dưới lên”, tức là từ phần tử có nghĩa nhỏ nhất, giá trị thuộc tính;
- tại giao diện của hệ thống quản lý dữ liệu cái mà không nằm trong hệ thống.

Các phiên bản hiện thời của các tiêu chuẩn thuộc chuỗi tiêu chuẩn về chất lượng dữ liệu cái trong bộ tiêu chuẩn TCVN 10249 (ISO 8000) bao gồm các yêu cầu có thể được kiểm tra bằng máy tính để trao đổi giữa các tổ chức và các hệ thống về dữ liệu cái, bao gồm dữ liệu đặc trưng. Các tiêu chuẩn xác định chất lượng của giá trị thuộc tính được trao đổi trong các thông điệp dữ liệu cái.

Tiêu chuẩn này miêu tả những điều cơ bản về chất lượng dữ liệu cái, định nghĩa các thuật ngữ có liên quan và các yêu cầu cụ thể về cả dữ liệu và tổ chức để có được chất lượng dữ liệu cái.

Tiêu chuẩn này bao gồm:

- đặc tả kỹ thuật về phạm vi của chuỗi tiêu chuẩn về chất lượng dữ liệu cái trong bộ tiêu chuẩn TCVN 10249 (ISO 8000);
- giới thiệu về dữ liệu cái;
- miêu tả về kiến trúc dữ liệu;
- tổng quan về nội dung mỗi phần trong chuỗi tiêu chuẩn.

Tiêu chuẩn này không bao gồm:

- các khía cạnh về chất lượng dữ liệu áp dụng cho tất cả loại dữ liệu ngoại trừ đó là dữ liệu cái;
- các khía cạnh về chất lượng dữ liệu áp dụng cho dữ liệu mà không phải là dữ liệu cái.

VÍ DỤ Dữ liệu giao dịch không phải là dữ liệu cái.

2 Tài liệu viện dẫn

Các tài liệu viện dẫn sau rất cần thiết cho việc áp dụng tiêu chuẩn này. Đối với các tài liệu viện dẫn ghi năm công bố thì áp dụng phiên bản được nêu. Đối với các tài liệu viện dẫn không ghi năm công bố thì áp dụng phiên bản mới nhất, bao gồm cả các sửa đổi, bổ sung (nếu có).

TCVN 10249-2 (ISO 8000-2)¹, *Chất lượng dữ liệu – Phần 2: Từ vựng*.

ISO/IEC 8824-1, Information technology – Abstract Syntax Notation One (ASN.1) – Part 1: Specification of basic notation (*Công nghệ thông tin – Cú pháp trừu tượng ký hiệu 1 (ASN.1) – Phần 1: Đặc tả cho ký hiệu cơ bản*).

3 Thuật ngữ và định nghĩa

Tiêu chuẩn này sử dụng các thuật ngữ và định nghĩa nêu trong TCVN 10249-2 (ISO 8000-2).

4 Thuật ngữ viết tắt

Tiêu chuẩn này sử dụng các thuật ngữ viết tắt sau:

ASN	Abstract Syntax Notation	Ký hiệu cú pháp trừu tượng
BOM	bill of materials	Danh mục vật tư
CPT	Current Procedural Terminology	Thuật ngữ dành cho các thủ tục hiện tại
DUNS	Data Universal Numbering System	Hệ thống số hóa dữ liệu quốc tế
EAN	European Article Number	Mã số mã vạch châu Âu
ERP	enterprise resource planning	Hoạch định tài nguyên doanh nghiệp
GTIN	Global Trade Identification Number	Mã số thương phẩm toàn cầu
HR	human resources	Nguồn nhân lực
id	identifier	Mã định danh
IPC	illustrated parts catalogue	Catalô thành phần minh họa
IUID	Item Unique Identification	Định danh hạng mục đơn nhất
MDM	master data message	Thông điệp dữ liệu cái
MDR	master data record	Bản ghi dữ liệu cái
MRO	maintenance, repair and overhaul	Bảo trì, sửa chữa và đại tu
sku	stockkeeping unit	đơn vị phân loại hàng hóa tồn kho
UCC	Uniform Code Council	Hội đồng mã hóa thống nhất
UML	Unified Modeling Language	Ngôn ngữ mô hình hóa thống nhất
UPC	Universal Product Code	Mã sản phẩm toàn cầu

5 Dữ liệu cái

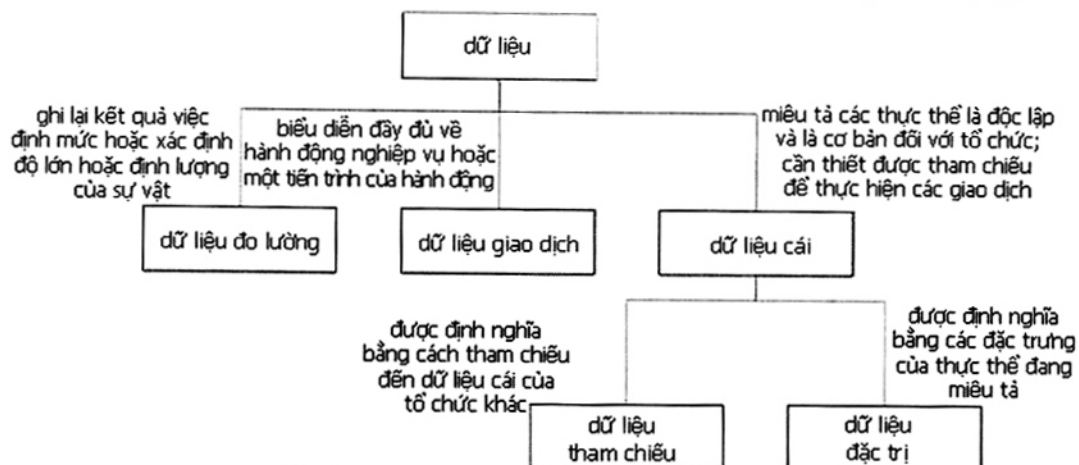
Bên trong tổ chức, dữ liệu cái được sử dụng để nhận diện và miêu tả những thứ có tầm quan trọng đối với tổ chức.

¹ ISO 8000-2:2012 thay thế cho ISO 8000-102:2009.

CHÚ THÍCH 1 Trong ứng dụng danh mục, dữ liệu cái được sử dụng để miêu tả những thứ được gọi là “hạng mục”.

Hình 1 mô tả cách phân loại dữ liệu, cho thấy đâu phù hợp dữ liệu cái.

CHÚ THÍCH 2 Hình 1 không dùng để phân loại hoàn toàn dữ liệu, ở đây chỉ nhằm cho biết khái niệm về dữ liệu cái.



Hình 1 – Phân loại dữ liệu (cho dữ liệu cái)

Dữ liệu cái điển hình được tham chiếu trong giao dịch thương mại thông qua mã định danh. Mã định danh là tham chiếu chung cho cả bản thân sự vật và cho cả bản ghi dữ liệu cái (MDR) có miêu tả về vật thể. MDR được xử lý chung tại kho lưu trữ trung tâm.

VÍ DỤ 1 Thông thường kho lưu trữ MDR trung tâm cho tổ chức thuộc về hệ thống hoạch định tài nguyên doanh nghiệp (ERP) hay hệ thống quản lý dữ liệu cái (MDM) của tổ chức.

CHÚ THÍCH 3 Một MDR đơn hợp lý nào có thể được biểu diễn bằng một số bản ghi vật lý trong hệ thống phần mềm.

VÍ DỤ 2 Trong việc thực thi cơ sở dữ liệu quan hệ, bản ghi dữ liệu cái có thể bao gồm các hàng từ nhiều bảng khác nhau.

CHÚ THÍCH 4 MDR miêu tả một số thứ có thể được nhận diện thông qua tham chiếu có sử dụng mã định danh. Một số thứ có thể được miêu tả bằng dữ liệu đặc trưng, được biểu diễn bằng dữ liệu thuộc tính. Ngoài ra, một số thứ có thể được miêu tả bằng chuỗi miêu tả hay định nghĩa.

Tham chiếu định danh là các tham chiếu được sử dụng để xử lý dữ liệu cái bằng những thứ khác.

VÍ DỤ 3 Mã định danh thuế thu nhập doanh nghiệp, số bảo hiểm quốc gia của cá nhân và số thành phần được gán bởi nhà sản xuất cho hạng mục sản xuất là tất cả ví dụ về tham chiếu định danh.

Để tham chiếu định danh có nghĩa, nó phải được gán bởi hệ thống định danh.

VÍ DỤ 4 Tổ chức khi đưa ra mã định danh có thể được quy định bằng siêu dữ liệu, điều này thông thường với mã định danh thuế nhưng số thành phần không có nghĩa nếu nhà sản xuất không công bố.

Miêu tả có thể là dữ liệu đặc trưng mà máy tính có thể hiểu được, trong đó được biểu diễn điển hình là giá trị thuộc tính hoặc dạng văn bản mà con người có thể đọc được. Một số thuộc tính là vi phân. Bởi vì như thế có thể dễ dàng xử lý chúng, các giá trị số học hoặc giá trị điều khiển hầu hết thường được sử dụng dạng vi phân.

Một trong những khía cạnh chính của việc quản lý chất lượng dữ liệu cái là quản lý bản sao. Cách tiếp cận phù hợp là quản lý và loại bỏ bản sao không thích hợp là phần quan trọng trong quản lý dữ liệu cái.

Đặc tính là vi phân khi xem xét bởi một tổ chức này và có thể là không vi phân khi xem xét bởi tổ chức khác.

Ví DỤ 5 Nhà sản xuất có thể có bản ghi dữ liệu cái sai khác với mỗi hạng mục sản xuất. Từ đánh giá của người mua, khi một số hạng mục sản xuất (được chế tạo bởi cùng nhà sản xuất hoặc khác nhà sản xuất) chia sẻ cùng các đặc tính phù hợp mẫu và chức năng, người mua có thể gộp nhóm theo hạng mục cung ứng đơn và gán "số kho hàng" như là tham chiếu định danh cho hạng mục cung ứng đó. Trong khi gộp nhóm một số hạng mục sản xuất thành một hạng mục cung ứng đơn, người mua đang đưa ra quyết định một hay nhiều đặc tính coi như là không vi phân mà nhà sản xuất coi là vi phân.

Đặc tính được coi là vi phân đối với một chức năng này bên trong tổ chức và có thể là không vi phân đối với chức năng khác bên trong cùng tổ chức đó.

Dữ liệu cái không nhất thiết cố định. Dĩ nhiên, số lượng đặc tính cần thiết để miêu tả sự vật sẽ đa dạng theo chức năng nghiệp vụ. Như sự đa dạng về số lượng đặc tính vi phân, MDR có thể bị vi phân khi các đặc tính được thêm vào hoặc bị thay đổi thành dạng vi phân. MDR có thể bị sao trùng khi các đặc tính bị loại bỏ hay bị thay đổi thành dạng không vi phân.

Ví dụ về dữ liệu cái bao gồm:

- dữ liệu cái bên bán: miêu tả điển hình thuật ngữ bên bán về địa chỉ và tình trạng hợp pháp. Rất nhiều dữ liệu cái bên bán bắt buộc được xác định theo luật pháp như là các yêu cầu chung cho công ty để có thể chỉ ra tất cả thực thể khi buôn bán.
- dữ liệu cái bên mua: miêu tả điển hình thuật ngữ khách hàng về thực thể buôn bán. Ít nhất nó cũng có thông tin liên hệ cần thiết để thực hiện trao đổi và có thể bao gồm thông tin bí mật như thông tin thẻ tín dụng.

CHÚ THÍCH 5 Nếu dữ liệu cá nhân được duy trì trong dữ liệu cái bên mua, đây có thể là chủ đề lập pháp bảo vệ dữ liệu như Đạo luật bảo vệ dữ liệu vương quốc Anh năm 1998.

- dữ liệu cái hạng mục hay vật tư: các dữ liệu cái điển hình này miêu tả các hạng mục hữu hình được trao đổi, kiểm kê hoặc mua sắm hợp lệ. Trong khi chúng thường bị hạn chế cho các hạng mục được mua sắm theo hợp đồng như vật tư sản xuất, chúng cũng có thể được sử dụng để cải tiến chất lượng cho việc phân tích chi phí có liên quan đến mua sắm bảo trì, sửa chữa và vận hành (MRO). Dữ liệu cái vật tư cũng có thể được sử dụng để hỗ trợ danh mục vật tư (BOM) hoặc cho việc thiết kế trong khi chúng có thể được tham chiếu đến catalô thành phần chung hoặc danh sách thành phần được xác định trước. Một số dữ liệu cái vật tư là catalô thành phần minh họa (IPC) hay danh sách thành phần thay thế máy.
- khái niệm hạng mục cung ứng: các dữ liệu cái này bao gồm tham chiếu đến dữ liệu cái hạng mục hoặc vật tư, cộng thêm thông tin đóng gói và định lượng.
- dữ liệu cái dịch vụ, thủ tục hoặc quy trình: các dữ liệu cái này vẫn không có liên quan về đảm bảo sức khỏe và nền công nghiệp sửa chữa xe vận tải trong đó hóa đơn tự động cho dịch vụ hay hoàn trả bảo hiểm được dùng chung. Điển hình là dịch vụ được miêu tả như là một thủ tục hoặc quy trình.

Ví DỤ 6 Mã thuật ngữ thủ tục hiện thời 4 (CPT-4) của Liên hiệp thuốc Hoa kỳ là ví dụ về dữ liệu cái thủ tục.

- dữ liệu cái tài sản: các dữ liệu cái này thông thường được sử dụng để trao đổi hạng mục có giá mua bán cao hơn giá trị tiền mặt định trước, hoặc có chi phí bị mất giá sau nhiều năm. Tài sản thông thường được gán với một mã định danh đơn nhất (số seri) và thường được gán với hạng mục có thể di chuyển trong đó ngày tháng (đôi khi là thời gian) và địa chỉ cần thiết để xác minh và

báo cáo. Việc mô hình hóa hiệu chỉnh cho dữ liệu cái tài sản là quan trọng để có thể trao đổi không chỉ địa chỉ và giá trị tài sản theo thời gian mà còn cho các hoạt động bảo trì và sửa chữa. Vấn đề điển hình với quản lý tài sản là thay đổi đặc tả kỹ thuật theo nhịp vòng đời tài sản. Việc quyết định tại tình trạng nào mà tài sản có thể bị sửa đổi cần yêu cầu việc tạo mới tài sản đã miêu tả thường là vấn đề cần thiết.

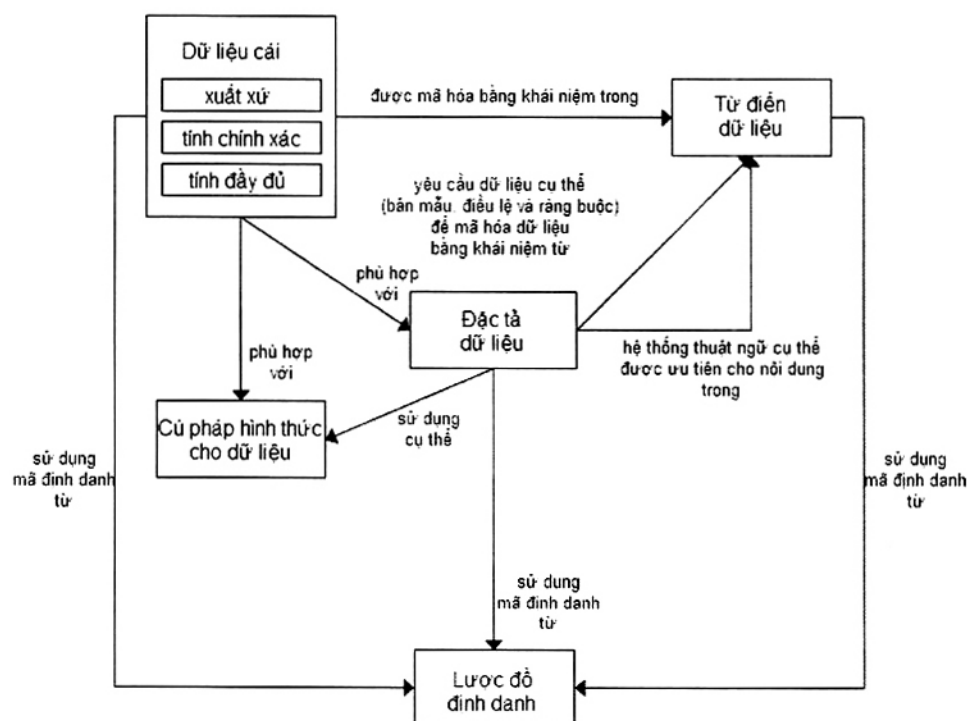
- dữ liệu cái vị trí: Khác với dịch vụ phân phối, ít khi xem dữ liệu cái địa chỉ tách biệt, nhưng khi tách biệt ra dữ liệu cái địa chỉ khỏi dữ liệu cái bên mua và bên bán thì nâng cao chất lượng dữ liệu. Mô hình dữ liệu cho dữ liệu cái địa chỉ là ví dụ cơ bản theo lý thuyết miêu tả địa chỉ vật lý tại đó tổng hợp vị trí chung nhất cung cấp tham chiếu chính xác. Trong thực tế, thông tin này cần bao gồm cả các phần miêu tả phân phối khác như địa chỉ thư tín.
- dữ liệu cái điểm liên lạc, dữ liệu cái nhân viên hoặc dữ liệu cái nguồn nhân lực (HR): các dữ liệu cái này cơ bản miêu tả riêng lẻ. Thông thường chúng bao gồm thông tin liên quan đến mối quan hệ với tổ chức tuyển dụng nhưng để xử lý tốt hơn thì dữ liệu giao dịch được thêm vào dữ liệu cái. Chúng thường chứa các thông tin bí mật.

CHÚ THÍCH 6 Dữ liệu bao gồm các dữ liệu cái này có thể được đặt luật lệ và là chủ đề lập pháp bảo vệ dữ liệu như Đạo luật bảo vệ dữ liệu vương quốc Anh năm 1998.

6 Kiến trúc dữ liệu cho dữ liệu cái

Điều này bao gồm kiến trúc mức cao của dữ liệu cái. Kiến trúc này có thể có khả năng ứng dụng hơn cả dữ liệu cái.

Hình 2 cho thấy kiến trúc dữ liệu cho dữ liệu cái.



Hình 2 – Kiến trúc dữ liệu cho dữ liệu cái

- Dữ liệu cái bao gồm thông tin về xuất xứ dữ liệu, độ chính xác dữ liệu và tính đầy đủ dữ liệu.
- Dữ liệu cái được mã hóa sử dụng các khái niệm trong từ điển dữ liệu.
- Dữ liệu cái phù hợp với đặc tả dữ liệu.
- Dữ liệu cái phù hợp với cú pháp hình thức.
- Đặc tả dữ liệu quy định các yêu cầu dữ liệu để mã hóa dữ liệu cái sử dụng các khái niệm trong từ điển dữ liệu.
- Đặc tả dữ liệu quy định việc sử dụng cú pháp hình thức.
- Dữ liệu cái, đặc tả dữ liệu và từ điển dữ liệu sử dụng mã định danh theo lược đồ định danh.

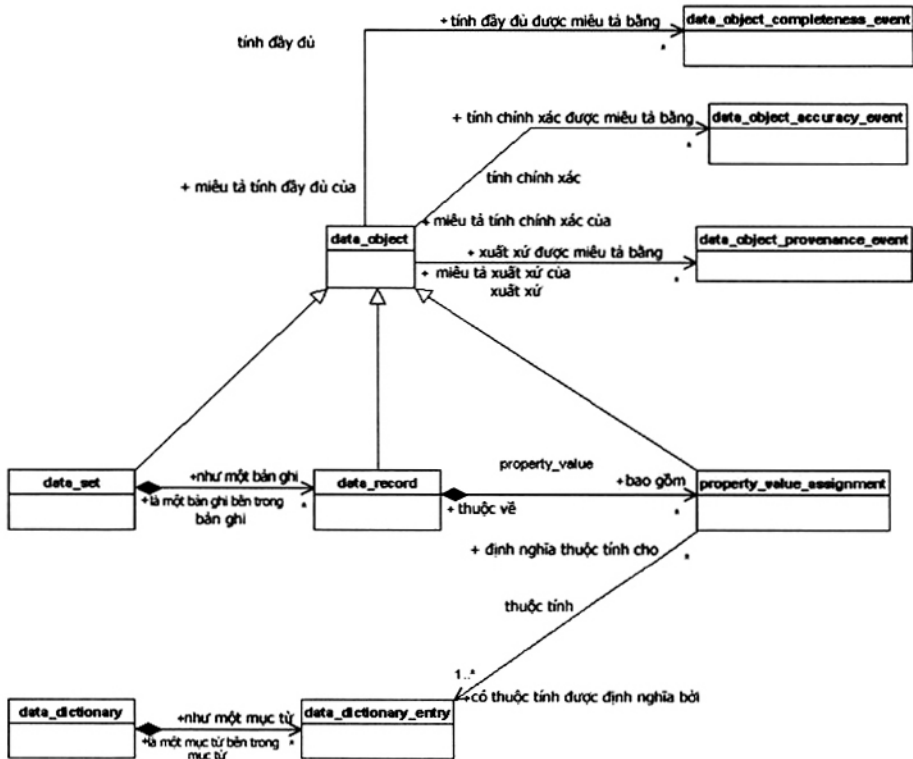
7 Mô hình dữ liệu mức cao

7.1 Tổng quan

Điều 7 bao gồm mô hình dữ liệu khái niệm mức cao của dữ liệu cái. Mục đích của mô hình dữ liệu này là cung cấp cái nhìn bao quát về cách thức các thực thể khóa trong mô hình dữ liệu tại TCVN 10249-110 (ISO 8000-110), TCVN 10249-120 (ISO/TS 8000-120), TCVN 10249-130 (ISO/TS 8000-130) và TCVN 10249-140 (ISO/TS 8000-140) phù hợp với nhau. Mô hình dữ liệu này có thể có khả năng ứng dụng nhiều hơn cả dữ liệu cái.

7.2 Biểu đồ

Biểu đồ lớp UML cho mô hình về mô hình dữ liệu mức cao được đưa ra trong Hình 3.



Hình 3 – Biểu đồ lớp UML cho mô hình dữ liệu lớp cao.

7.3 Thực thể

7.3.1 **data_dictionary** (từ điển dữ liệu)

data_dictionary (từ điển dữ liệu) là tổ hợp các đối tượng **data_dictionary_entry** (mục từ từ điển dữ liệu) cho phép tìm kiếm bằng mã định danh thực thể.

Định nghĩa thuộc tính:

thực thể: **data_dictionary_entry** (mục từ từ điển dữ liệu) cấu tạo nên **data_dictionary** (từ điển dữ liệu).

Xác nhận:

Mỗi **data_dictionary** (từ điển dữ liệu) có kiểu đối tượng không, đối tượng một mục từ hoặc nhiều đối tượng **data_dictionary_entry** (mục từ từ điển dữ liệu). Mỗi **data_dictionary_entry** (mục từ từ điển dữ liệu) là mục từ nằm trong một **data_dictionary** (từ điển dữ liệu) cụ thể.

7.3.2 **data_dictionary_entry** (mục từ từ điển dữ liệu)

data_dictionary_entry (mục từ từ điển dữ liệu) là miêu tả cho một mục từ đang chứa, ít nhất một mã định danh rõ ràng, một thuật ngữ và một định nghĩa.

Xác nhận:

Mỗi **data_dictionary_entry** (mục từ từ điển dữ liệu) là mục từ nằm trong một **data_dictionary** (từ điển dữ liệu) cụ thể. Mỗi **data_dictionary** (từ điển dữ liệu) có đối tượng mục từ không, đối tượng một mục từ hoặc nhiều đối tượng **data_dictionary_entry** (mục từ từ điển dữ liệu).

Mỗi **data_dictionary_entry** (mục từ từ điển dữ liệu) định nghĩa thuộc tính cho đối tượng không, đối tượng một và nhiều đối tượng **property_value_assignment** (gán giá trị thuộc tính).

Mỗi **property_value_assignment** (gán giá trị thuộc tính) có thuộc tính được định nghĩa bằng một hoặc nhiều đối tượng **data_dictionary_entry** (mục từ từ điển dữ liệu).

7.3.3 **data_record** (bản ghi dữ liệu)

data_record (bản ghi dữ liệu) là **data_object** (đối tượng dữ liệu) là tập các đối tượng **property_value_assignment** (gán giá trị thuộc tính).

CHÚ THÍCH Thuộc tính có thể mang tính ẩn theo cặp giá trị thuộc tính.

Định nghĩa thuộc tính:

property_value (giá trị thuộc tính): **property_value_assignment** (gán giá trị thuộc tính) cấu tạo nên **data_record** (bản ghi dữ liệu).

Xác nhận:

Mỗi **data_record** (bản ghi dữ liệu) có chứa đối tượng không, đối tượng một hoặc nhiều đối tượng **property_value_assignment** (gán giá trị thuộc tính). Mỗi **property_value_assignment** (gán giá trị thuộc tính) thuộc về một **data_record** (bản ghi dữ liệu) cụ thể.

Mỗi **data_record** (bản ghi dữ liệu) là bản ghi nằm trong một **data_set** (tập dữ liệu) cụ thể. Mỗi **data_set** (tập dữ liệu) có đối tượng bản ghi không, đối tượng một bản ghi hoặc nhiều đối tượng **data_record** (bản ghi dữ liệu).

7.3.4 data_set (tập dữ liệu)

data_set (tập dữ liệu) là **data_object** (đối tượng dữ liệu) là tập các đối tượng **data_record** (bản ghi dữ liệu), trong đó có thể có thứ tự hoặc không.

Định nghĩa thuộc tính:

record (bản ghi): **data_record** (bản ghi dữ liệu) tạo ra **data_set** (tập dữ liệu).

Xác nhận:

Mỗi **data_set** (tập dữ liệu) có đối tượng bản ghi không, đối tượng một hoặc nhiều **data_record** (bản ghi dữ liệu). Mỗi **data_record** (bản ghi dữ liệu) là bản ghi thuộc về một **data_set** (tập dữ liệu) cụ thể.

7.3.5 data_object (đối tượng dữ liệu)

data_object (đối tượng dữ liệu) là bất cứ thứ gì được sử dụng để biểu thị một số thứ khác.

Định nghĩa thuộc tính:

accuracy (độ chính xác): **data_object_accuracy_event** (sự kiện độ chính xác đối tượng dữ liệu) cung cấp thông tin về độ chính xác của **data_object** (đối tượng dữ liệu).

completeness (tính đầy đủ): **data_object_completeness_event** (sự kiện tính đầy đủ đối tượng dữ liệu) cung cấp thông tin về tính đầy đủ của **data_object** (đối tượng dữ liệu).

provenance (xuất xứ): **data_object_provenance_event** (sự kiện xuất xứ đối tượng dữ liệu) cung cấp thông tin về xuất xứ của **data_object** (đối tượng dữ liệu).

Xác nhận:

Mỗi **data_object** (đối tượng dữ liệu) có xuất xứ được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_provenance_event** (sự kiện xuất xứ đối tượng dữ liệu). Mỗi **data_object_provenance_event** (sự kiện xuất xứ đối tượng dữ liệu) miêu tả xuất xứ của một **data_object** (đối tượng dữ liệu) cụ thể.

Mỗi **data_object** (đối tượng dữ liệu) có độ chính xác được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_accuracy_event** (sự kiện độ chính xác đối tượng dữ liệu). Mỗi **data_object_accuracy_event** (sự kiện độ chính xác đối tượng dữ liệu) miêu tả độ chính xác của một **data_object** (đối tượng dữ liệu) cụ thể.

Mỗi **data_object** (đối tượng dữ liệu) có tính đầy đủ được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_completeness_event** (sự kiện tính đầy đủ đối tượng dữ liệu). Mỗi **data_object_completeness_event** (sự kiện tính đầy đủ đối tượng dữ liệu) miêu tả tính đầy đủ của một **data_object** (đối tượng dữ liệu) cụ thể.

7.3.6 data_object_accuracy_event (sự kiện độ chính xác đối tượng dữ liệu)

data_object_accuracy_event (sự kiện độ chính xác đối tượng dữ liệu) là sự kiện trong đó ghi lại thông tin độ chính xác dữ liệu.

Xác nhận:

Mỗi **data_object_accuracy_event** (sự kiện độ chính xác đối tượng dữ liệu) miêu tả độ chính xác của một **data_object** (đối tượng dữ liệu) cụ thể.

Mỗi **data_object** (đối tượng dữ liệu) có độ chính xác được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_accuracy_event** (sự kiện độ chính xác đối tượng dữ liệu).

7.3.7 data_object_completeness_event (sự kiện tính đầy đủ đối tượng dữ liệu)

data_object_completeness_event (sự kiện tính đầy đủ đối tượng dữ liệu) là sự kiện trong đó ghi lại thông tin tính đầy đủ dữ liệu.

Xác nhận:

Mỗi **data_object_completeness_event** (sự kiện tính đầy đủ đối tượng dữ liệu) miêu tả tính đầy đủ của một **data_object** (đối tượng dữ liệu) cụ thể.

Mỗi **data_object** (đối tượng dữ liệu) có tính đầy đủ được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_completeness_event** (sự kiện tính đầy đủ đối tượng dữ liệu).

7.3.8 data_object_provenance_event (sự kiện xuất xứ đối tượng dữ liệu)

data_object_provenance_event (sự kiện xuất xứ đối tượng dữ liệu) là sự kiện trong đó ghi lại thông tin xuất xứ dữ liệu.

Xác nhận:

Mỗi **data_object_provenance_event** (sự kiện xuất xứ đối tượng dữ liệu) miêu tả xuất xứ của một **data_object** (đối tượng dữ liệu) cụ thể. Mỗi **data_object** (đối tượng dữ liệu) có xuất xứ được miêu tả bằng đối tượng không, đối tượng một hoặc nhiều đối tượng **data_object_provenance_event** (sự kiện xuất xứ đối tượng dữ liệu).

7.3.9 property_value_assignment (gán giá trị thuộc tính)

property_value_assignment (gán giá trị thuộc tính) là **data_object** (đối tượng dữ liệu) là một cặp gồm giá trị và mã định danh của thuộc tính được định nghĩa trong từ điển dữ liệu.

Định nghĩa thuộc tính:

ID: chuỗi có chỉ ra rõ ràng **property_value_assignment** (gán giá trị thuộc tính) bên trong tổ chức đã tạo ra nó.

CHÚ THÍCH 1 Mã định danh chỉ cần là đơn nhất hoặc có nghĩa bên trong tổ chức đã tạo **property_value_assignment** (gán giá trị thuộc tính).

CHÚ THÍCH 2 Định dạng cho mã định danh không được quy định trong tiêu chuẩn này.

property (thuộc tính) **data_dictionary_entry** (mục từ từ điển dữ liệu) định nghĩa thuộc tính cho giá trị đã gán.

Xác nhận:

Mỗi **property_value_assignment** (gán giá trị thuộc tính) thuộc về một **data_record** (bản ghi dữ liệu) cụ thể. Mỗi **data_record** (bản ghi dữ liệu) có chứa các đối tượng không, đối tượng một hoặc nhiều **property_value_assignment** (gán giá trị thuộc tính).

Mỗi **property_value_assignment** (gán giá trị thuộc tính) có thuộc tính được định nghĩa bằng một hoặc nhiều đối tượng **data_dictionary_entry** (mục từ từ điển dữ liệu). Mỗi **data_dictionary_entry** (mục từ từ điển dữ liệu) định nghĩa thuộc tính cho đối tượng không, đối tượng một và nhiều đối tượng **property_value_assignment** (gán giá trị thuộc tính).

8 Tổng quan về bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái

Bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái bao gồm các phần sau với tiêu đề chung là *Chất lượng dữ liệu – Dữ liệu cái*:

- Phần 100, *Trao đổi dữ liệu đặc trưng: Tổng quan*;
- Phần 110, *Trao đổi dữ liệu đặc trưng: Cú pháp, mã hóa ngữ nghĩa và sự phù hợp với đặc tả dữ liệu*;
- Phần 120, *Trao đổi dữ liệu đặc trưng: Xuất xứ*;
- Phần 130, *Trao đổi dữ liệu đặc trưng: Độ chính xác*;
- Phần 140, *Trao đổi dữ liệu đặc trưng: Tính đầy đủ*.

Phần 100 bao gồm:

- Phạm vi của bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái;
- Giới thiệu về dữ liệu cái;
- Kiến trúc dữ liệu của bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái;
- Tổng quan về khái niệm cho các phần tiêu chuẩn khác.

Phần 2² là phần quy định từ vựng chung và có phần từ vựng được sử dụng trong bộ tiêu chuẩn TCVN 10249 (ISO 8000) về chất lượng dữ liệu cái.

CHÚ THÍCH 1 Phần 2 chỉ bao gồm các thuật ngữ, định nghĩa và từ viết tắt được sử dụng trong các phần tiêu chuẩn về chất lượng dữ liệu cái đã công bố. Do đó không bao gồm các từ vựng toàn bộ về chất lượng dữ liệu cái.

Phần 110 đến 140 bao gồm các yêu cầu có thể được kiểm tra bằng máy tính để trao đổi dữ liệu cái, giữa tổ chức và hệ thống, bao gồm dữ liệu đặc trưng. Các tiêu chuẩn này chỉ ra chất lượng của giá trị thuộc tính có thể được trao đổi trong các thông điệp dữ liệu cái.

Phần 110 bao gồm:

- yêu cầu cho thông điệp dữ liệu cái, các yêu cầu này được sử dụng trong trao đổi hạng mục dữ liệu đặc trưng:
 - sự dính dáng đến cú pháp hình thức;
 - mã hóa ngữ nghĩa;

CHÚ THÍCH 2 Mã hóa ngữ nghĩa là kỹ thuật thay thế các thuật ngữ thuộc ngôn ngữ tự nhiên có trong thông điệp bằng mã định danh tham chiếu đến các mục từ từ điển dữ liệu.

- sự phù hợp với đặc tả dữ liệu bên mua;
- mô hình nghiệp vụ

Phần 120 là phần bổ sung cho phần 110 bao gồm:

- yêu cầu để thu thập và trao đổi thông tin xuất xứ dữ liệu;
- mô hình dữ liệu cho thông tin xuất xứ dữ liệu.

² ISO 8000-2:2012 thay thế cho ISO 8000-102:2009

Phần 130 là phần bổ sung cho phần 120 bao gồm:

- yêu cầu để thu thập và trao đổi thông tin độ chính xác dữ liệu theo mẫu biểu diễn và cảnh báo về độ chính xác dữ liệu;
- khái niệm mô hình dữ liệu cho thông tin độ chính xác dữ liệu theo mẫu biểu diễn và cảnh báo về độ chính xác dữ liệu.

Phần 140 là phần bổ sung cho phần 120 bao gồm:

- yêu cầu để thu thập và trao đổi thông tin tính đầy đủ dữ liệu theo mẫu biểu diễn và cảnh báo về tính đầy đủ dữ liệu;
- khái niệm mô hình dữ liệu cho thông tin tính đầy đủ dữ liệu theo mẫu biểu diễn và cảnh báo về tính đầy đủ dữ liệu;

Phụ lục A

(quy định)

Nhận diện tài liệu

Để cung cấp cho việc nhận diện minh bạch một đối tượng thông tin trong hệ thống mở, định danh đối tượng

{tiêu chuẩn ISO 8000 phần (100) phiên bản (1)}

được gán cho tiêu chuẩn này. Như thế giá trị được định nghĩa trong ISO/IEC 8824-1, và được miêu tả trong ISO 10303-1.

Phụ lục B

(tham khảo)

Danh mục hạng mục

B.1 Tổng quát

Mã định danh thực hiện theo luật quyết định trong quản lý chuỗi cung ứng và hỗ trợ vòng đời sản phẩm. Ở đây có ba danh mục hạng mục, mỗi danh mục có kiểu mã định danh riêng:

- đối tượng vật lý: mã hiệu phân chia tài sản và số seri;
- hạng mục sản xuất: mã hiệu cấu kiện hoặc mẫu mã;
- hạng mục cung ứng: mã hiệu lô.

B.2 Đối tượng vật lý

B.2.1 Mã hiệu phân chia tài sản hoặc số seri

Mã hiệu phân chia tài sản hoặc số seri là số đơn nhất cho một đối tượng vật lý đơn. Nếu hạng mục có giá trị hay cảnh báo giảm, nếu đó là thuế, hoặc cần phân chia, hạng mục phải có mã hiệu phân chia tài sản hoặc số seri. Trong khi hầu hết thường dùng số seri, chúng lại thường không được gán bởi nhà sản xuất gốc. Có một số lược đồ phân công và lược đồ ủy nhiệm để đảm bảo rằng số seri là đơn nhất trong miền hoặc khoảng thời gian có đảm bảo rằng không có hai hạng mục dùng chung mã hiệu giống nhau. Nếu số seri không được gán bởi nhà sản xuất thì bên cung ứng hoặc bên mua phải thêm vào mã hiệu phân chia tài sản. Số seri thông thường được tạo như một chuỗi số liên tục hoặc mã định danh chữ-số và tổ chức phân chia phải duy trì cơ sở dữ liệu về gốc phân chia chung và chủ sở hữu hiện thời của hạng mục.

Số seri thông thường được tạo như một chuỗi số liên tục hoặc mã định danh chữ-số và tổ chức phân chia phải duy trì cơ sở dữ liệu về gốc phân chia chung và chủ sở hữu hiện thời của hạng mục.

B.2.2 Mã hiệu lô

Một biến thể của số seri là mã hiệu lô. Mặc dù mã hiệu lô có liên kết tới một nhóm hạng mục hoặc một số tài sản hơn là một trường hợp hạng mục, mã hiệu lô được sử dụng như số seri cho mục đích kiểm soát chất lượng và cảnh báo chất lượng. Mã hiệu lô thông thường được tìm thấy trên hạng mục hao mòn như thức ăn hoặc hóa chất hoặc các hạng mục dễ tiêu hủy để cho mục đích đảm bảo chất lượng và truy tìm xuất xứ. Định danh hạng mục đơn nhất (IUID) là ví dụ về mã hiệu tài sản chung.

B.3 Hạng mục sản xuất

B.3.1 Mã hiệu sản phẩm, thành phần và mẫu mã

Mã hiệu sản phẩm chỉ rõ kiểu hạng mục hoặc tài sản mà nhà sản xuất chế tạo, đã chế tạo hoặc hoạch định chế tạo. Mã hiệu cấu kiện hoặc mã hiệu mẫu mã là kiểu mã hiệu sản phẩm mà nhà sản xuất hoặc bên cung ứng sử dụng để chỉ rõ nhóm hạng mục phân phối là có thể thay thế trong ứng dụng cụ thể. Mục đích chính của mã hiệu sản phẩm, thành phần hoặc mẫu mã là hỗ trợ bán hàng và giới thiệu sản phẩm, các mã hiệu này là các số thường được dùng để sắp xếp lại. Nhà sản xuất và bên cung ứng

thường sử dụng các mã hiệu sản phẩm, thành phần hoặc mẫu mã khác nhau để chỉ rõ tài sản hoặc hạng mục được thiết kế để di chuyển thông qua các kênh phân phối khác nhau. Rõ ràng đây là việc thông dụng trong nền công nghiệp điện tử bán lẻ (ví dụ nhận diện các mẫu mã tivi từ cùng một nhà sản xuất phải có mã hiệu mẫu mã khác nhau phụ thuộc vào đại lý bán lẻ là nơi mà chúng được bán). Mã hiệu này thường được sử dụng trong nền công nghiệp hóa dầu, tại đó cùng một sản phẩm được bán theo nhiều nhánh khác nhau theo nhiều cửa hàng khác nhau.

Mã hiệu cấu kiện thường được thiết kế có bao gồm một số mẫu phân loại và thường bao gồm các đặc tính đã mã hóa của hạng mục. Trong khi đó, mã hiệu cấu kiện không cần đơn nhất, mã hiệu này thường có không hiếm công ty sử dụng mã số thành phần hoặc mẫu mã cho loại hàng.

Nhiều sáng kiến được thiết kế để tạo ra mã hiệu cấu kiện toàn cầu. Hầu hết đều cần có thêm một tiền tố để chỉ ra đơn nhất nhà sản xuất hoặc bên cung ứng là nơi phát ra mã hiệu. Hầu hết là mã vạch như mã sản phẩm toàn cầu (UPC) hoặc thay thế của nó, mã số thương phẩm toàn cầu (GTIN) phát ra bởi GS1, trước đây là hợp nhất giữa Hội đồng mã hóa thống nhất (UCC) Mỹ và Tổ chức mã số mã vạch Châu Âu (EAN). Nội dung cơ bản của mã hiệu UCC/EAN UPC và GTIN là kiểm soát tập trung các tiền tố nhà sản xuất hoặc bên cung ứng đơn nhất có liên quan đến sự hiểu biết rằng nhà sản xuất hay bên cung ứng được kiểm soát hậu tố phải là đơn nhất cho nhà sản xuất hay bên cung ứng đó. Các số khác như chỉ ra cùng tiền tố của mã hiệu cấu kiện là mã số định danh kinh doanh thương mại như mã hiệu hệ thống số hóa dữ liệu quốc tế (DUNS).

Mặc dù nhiều hạng mục có thể có chung mã hiệu sản phẩm hoặc mẫu mã, mỗi số seri phải là đơn nhất cho một hạng mục. Tồn tại mối quan hệ một-nhiều giữa mã hiệu sản phẩm và số seri (hạng mục có chung mã hiệu sản phẩm nhưng phải khác số seri).

B.3.2 Thuộc tính của hạng mục sản xuất

Thuộc tính của hạng mục sản xuất bao gồm ba loại sau:

- số tham chiếu như mã hiệu cấu kiện, mã hiệu mẫu mã hoặc mã thủ tục. Các tham chiếu này nhận diện đơn nhất một hạng mục xét theo khía cạnh nhà sản xuất hoặc bên cung ứng và hầu hết thường có liên quan đến đặc tả kỹ thuật;
- thuộc tính động như giá cả hay tính sẵn sàng (bao gồm địa điểm);
- thuộc tính thuộc tự nhiên chủ quan là không phù hợp để người mua sử dụng cho mục đích phân biệt khách quan về sự thích hợp, kiểu mẫu hoặc chức năng của hạng mục.

B.4 Hạng mục cung ứng

B.4.1 Mã hiệu kho

Mã hiệu kho, mã hiệu kiểm kê hay đơn vị phân loại hàng hóa tồn kho là số được phát hành bởi người mua không chỉ để phân chia kiểm kê nội bộ mà còn là phương pháp chính hỗ trợ tương tác và nguồn cạnh tranh. Như vậy nhà sản xuất và bên cung ứng cũng như người mua, họ cũng gán mã hiệu kho cho cái họ mua, lắp ráp hoặc chế tạo, và trong mã hiệu nhà kho thường trở thành mã hiệu cấu kiện khi đưa ra ngoài.

Hầu hết hệ thống quản lý kiểm kê phải sử dụng mã hiệu kho để kết nối nhiều nguồn với nhau để cung ứng, hoặc nhiều nhà cung ứng cho cùng mã hiệu cấu kiện nhà sản xuất hoặc nhà sản xuất thay thế và mã hiệu cấu kiện. Đây là khái niệm thay thế để phân biệt mã hiệu kho và mã hiệu cấu kiện.

B.4.2 Thuộc tính của hạng mục cung ứng

Thuộc tính của hạng mục cung ứng biểu diễn các đặc tính khác nhau của hạng mục. Bao gồm các thuộc tính miêu tả sự thích hợp, kiểu mẫu hoặc chức năng của hạng mục và cho phép hạng mục này phân biệt với hạng mục khác. Thuộc tính phân biệt là đơn vị chung khi đo đạc hoặc giá trị đã đếm được mặc dù nó có thể ở dạng văn bản tự do cấu thành một thuộc tính phân biệt.

VÍ DỤ Văn bản trong biển báo đường bộ.

Thư mục tài liệu tham khảo

- [1] TCVN 8244-2:2010 (ISO 3534-2:2006), Thống kê học – Từ vựng và ký hiệu – Phần 2: Thống kê và ứng dụng.
 - [2] TCVN 10249-1 (ISO/TS 8000-1), *Chất lượng dữ liệu – Phần 1: Tổng quan.*
 - [3] TCVN 10249-110 (ISO 8000-110), *Chất lượng dữ liệu – Phần 110: Dữ liệu cái – Trao đổi dữ liệu đặc trưng - Cú pháp, mã hóa ngữ nghĩa và sự phù hợp với đặc tả dữ liệu.*
 - [4] TCVN 10249-120 (ISO/TS 8000-120), *Chất lượng dữ liệu – Phần 120: Dữ liệu cái – Trao đổi dữ liệu đặc trưng - Xuất xứ.*
 - [5] TCVN 10249-130 (ISO/TS 8000-130), *Chất lượng dữ liệu – Phần 130: Dữ liệu cái – Trao đổi dữ liệu đặc trưng - Độ chính xác.*
 - [6] TCVN 10249-140 (ISO/TS 8000-140), *Chất lượng dữ liệu – Phần 140: Dữ liệu cái – Trao đổi dữ liệu đặc trưng - Tính đầy đủ.*
 - [7] ISO 10303-1, *Industrial automation systems and integration – Product data representation and exchange – Part 1: Overview and fundamental principles.*
-