

TCVN 7789- 2 : 2007
ISO/IEC 11179-2 : 2004

Xuất bản lần 1

CÔNG NGHỆ THÔNG TIN – SỔ ĐĂNG KÝ
SIÊU DỮ LIỆU(MDR) – PHẦN 2: PHÂN LOẠI
Information technology — Metadata registries (MDR) — Part 2: Classification

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 Thuộc tính của lược đồ phân loại.....	8
5 Cơ cấu phân loại một lần trong mỗi mục được quản trị.....	11
Phụ lục A.....	12
Tài liệu tham khảo.....	16

Lời nói đầu

TCVN 7789-2 : 2007 hoàn toàn tương đương với **ISO/IEC 11179-2 : 2004**

TCVN 7789-2 : 2007 do Ban kỹ thuật tiêu chuẩn TCVN/TC 154 "*Quá trình, các yếu tố dữ liệu và tài liệu thương mại, công nghiệp và hành chính*" biên soạn, Tổng cục Tiêu chuẩn Đo lường Chất lượng đề nghị, Bộ Khoa học và Công nghệ công bố.

Công nghệ thông tin - Sổ đăng ký siêu dữ liệu (MDR) Phần 2: Phân loại

Information technology — Metadata registries (MDR) - Part 2: Classification

1 Phạm vi áp dụng

Tiêu chuẩn này trình bày lại và bổ sung thêm các thủ tục và kỹ thuật đăng ký các lược đồ phân loại trong TCVN 7789-3:2007 và phân loại các mục được quản trị trong một MDR. Có thể phân loại tất cả các kiểu mục được quản trị, bao gồm các lớp đối tượng, các đặc tính, các cách biểu diễn, các miền giá trị và các khái niệm phân tử dữ liệu cũng như bản thân các phân tử dữ liệu.

Tiêu chuẩn này xây dựng một tập các nguyên tắc, phương pháp và thủ tục quy định những yêu cầu cần thiết (tối thiểu) để tài liệu hóa liên kết giữa các kiểu mục được quản trị khác nhau với một hoặc nhiều lược đồ phân loại. Tiêu chuẩn này bao gồm các tên, các định nghĩa và các khía cạnh khác của lược đồ phân loại và nội dung của nó. Có thể sao chép giữ lại các lược đồ phân loại và nội dung của chúng thông qua việc sử dụng một tập các thuộc tính. Các thuộc tính riêng trong tiêu chuẩn phù hợp với cấu trúc nội dung của các thuộc tính này. Người sử dụng có thể mở rộng tập các thuộc tính khi cần thiết. Thông tin bổ sung có thể trợ giúp thêm cho nguyên tắc phân loại hoặc bản thể học. Ví dụ; để cung cấp một bộ các hạn định đề xuất có thể áp dụng cho lớp đối tượng đó, thì đặc tính, hoặc quy định về trình bày để hạn định một cách đầy đủ hơn việc phân loại mục được quản trị riêng. Tiêu chuẩn này tóm tắt các thuộc tính cơ sở và mô hình được quy định trong TCVN 7789-3:2007.

Ví dụ trong tiêu chuẩn này chỉ ra cách lựa chọn các thành phần của phân tử dữ liệu có thể liên kết với lược đồ phân loại qua các thuộc tính được quy định trong tiêu chuẩn. Việc sử dụng một hoặc nhiều lược đồ phân loại để cung cấp một cơ sở khái niệm cho việc phát triển của siêu dữ liệu sẽ nâng cao về tính trong sáng của ngữ nghĩa và tính toàn vẹn của thiết kế.

Tiêu chuẩn này không thiết lập một lược đồ phân loại đặc biệt ưu việt. Sự thừa nhận cách tiếp cận phân loại và/hoặc nhận thức luận riêng nằm ngoài phạm vi tiêu chuẩn này. Các cách tiếp cận đó được đưa ra bởi các ban kỹ thuật tiêu chuẩn khác và/hoặc hoàn toàn thích hợp với một miền ngôn ngữ riêng. Khả năng của lược đồ phân loại và tiện ích của nội dung thích hợp đối với các lĩnh vực cạnh

TCVN 7789-2 : 2007

tranh. Các ban kỹ thuật tiêu chuẩn khác đã hoặc đang xây dựng các ngôn ngữ quy định để sử dụng trong việc phân loại và/hoặc các kỹ thuật và cấu trúc riêng có thể được hỗ trợ bởi tiêu chuẩn này. Ví dụ, tổ chức tiêu chuẩn thông tin quốc gia (NISO) xây dựng tiêu chuẩn cho việc phát triển từ điển chuyên đề. Điều này thích hợp với mỗi cấu trúc phân loại tài liệu hóa như cách thức xây dựng và mở rộng và duy trì nó. Các thuộc tính này có thể được bổ sung bởi nguyên tắc mở rộng cho các thuộc tính được quy định trong tiêu chuẩn này. Mặc dù các nguyên tắc mở rộng không có trong tiêu chuẩn này, nhưng vẫn được đề cập ở đây.

Mỗi tổ chức có thẩm quyền đăng ký, được mô tả và quy định trong TCVN 7789-6, có thể phân loại mục được quản trị theo các lược đồ phân loại, cấu trúc và nội dung thích hợp. Trong việc tài liệu hóa các khía cạnh về phân loại các mục được quản trị, tổ chức có thẩm quyền đăng ký có thể sử dụng các nguyên tắc, phương pháp, thủ tục và thuộc tính được quy định trong tiêu chuẩn này.

2 Tài liệu viện dẫn

TCVN 7789 (tất cả các phần), Công nghệ thông tin — Sổ đăng ký siêu dữ liệu (MDR)

3 Thuật ngữ và định nghĩa

Trong tiêu chuẩn này áp dụng các thuật ngữ và định nghĩa sau đây:

3.1

Thuộc tính (attribute)

Đặc điểm của một đối tượng hoặc một thực thể.

3.2

Lược đồ phân loại (classification scheme)

Thông tin mô tả về việc sắp xếp hoặc phân chia các đối tượng vào các nhóm dựa trên cơ sở các đặc điểm chung của các đối tượng đó.

3.3

Mục lược đồ phân loại (classification scheme item)

CSI

Mục nội dung trong lược đồ phân loại.

3.4

Quan hệ mục lược đồ phân loại (classification scheme item relationship)

Quan hệ giữa các mục lược đồ phân loại.

3.5

Mô tả kiểu quan hệ của mục lược đồ phân loại (classification scheme item relationship type description)

Mô tả kiểu quan hệ giữa một mục lược đồ phân loại với một hoặc nhiều mục lược đồ phân loại khác trong một lược đồ phân loại.

3.6

Tên kiểu của mục lược đồ phân loại (classification scheme item type name)

Tên kiểu của mục lược đồ phân loại.

3.7

Giá trị mục lược đồ phân loại (classification scheme item value)

Trường hợp cụ thể của một mục lược đồ phân loại.

3.8

Thành phần lược đồ phân loại (classification scheme membership)

Quan hệ của lược đồ phân loại với các mục của nó.

3.9

Tên kiểu lược đồ phân loại (classification scheme type name)

Tên gọi của kiểu lược đồ phân loại.

3.10

Khái niệm (concept)

Đơn vị tri thức được tạo bởi một kết hợp duy nhất của các đặc điểm.

3.11

Tên (name)

Phương tiện ban đầu định danh các đối tượng và khái niệm của con người.

3.12

Mục nhập thuật ngữ (terminological entry)

Mục nhập bao gồm thông tin về các đơn vị thuật ngữ cho một mục được quản trị cụ thể trong ngữ cảnh (phạm vi chủ thể)

4 Thuộc tính của lược đồ phân loại

Lược đồ phân loại phải được đăng ký trong một MDR bằng việc ghi lại các thuộc tính của chúng. Tối thiểu, trong một lược đồ phân loại được đăng ký phải có một bản ghi quản trị và một tên kiểu lược đồ phân loại. Các thuộc tính khác được liệt kê trong mục này có thể được ghi lại một cách tùy ý.

Bảng sau đây liệt kê các thuộc tính của một hệ thống phân loại có thể được ghi trong một MDR. Hình vẽ trong phụ lục A liên quan đến sơ đồ UML có thuộc tính xuất hiện. Các thuộc tính in **đậm** là các kiểu dữ liệu gồm nhiều thành phần thuộc tính.

Thuộc tính	Lần xuất hiện	Hình
<i>Designation-name</i> (ký hiệu quy ước-tên)	Một lần trong mỗi đoạn ngôn ngữ của mục nhập thuật ngữ	A.2
<i>Designation-preferred designation</i> (ký hiệu quy ước -hệ ký hiệu ưu tiên)	Không hoặc một lần trong mỗi đoạn ngôn ngữ của mục nhập thuật ngữ	A.2
<i>Designation-language identifier</i> (ký hiệu quy ước –thẻ định danh ngôn ngữ)	Một lần trong mỗi đoạn ngôn ngữ mục nhập thuật ngữ	A.2
<i>Definition-definition text</i> (định nghĩa–văn bản định nghĩa)	Một lần trong mỗi đoạn ngôn ngữ của mục nhập thuật ngữ	A.2
<i>Definition-preferred definition</i> (định nghĩa–định nghĩa ưu tiên)	Không hoặc một lần trong mỗi đoạn ngôn ngữ của mục nhập thuật ngữ	A.2
<i>Definition-source reference</i> (định nghĩa–tham chiếu nguồn)	Không hoặc một lần trong mỗi đoạn ngôn ngữ của mục nhập thuật ngữ	A.2

<i>Definition-language identifier</i> (định nghĩa–thẻ định danh ngôn ngữ)	Một lần trong mỗi đoạn ngôn ngữ trong mỗi Mục nhập thuật ngữ	A.2
Context-administration record (ngữ cảnh–bản ghi quản trị)	Một lần trong mỗi ngữ cảnh	A.2
<i>Context – description</i> (ngữ cảnh–mô tả)	Một lần trong mỗi ngữ cảnh	A.2
<i>Context – description identifier</i> (ngữ cảnh–thẻ định danh mô tả)	Không hoặc một lần trong mỗi ngữ cảnh	A.2
<i>Classification Scheme – type name</i> (lược đồ phân loại– kiểu tên)	Một lần trong mỗi lược đồ phân loại	A.1
<i>Classification Scheme Item – value</i> (mục lược đồ phân loại – giá trị)	Một lần trong mỗi mục lược đồ phân loại	A.1
<i>Classification Scheme Item – type name</i> (mục lược đồ phân loại – kiểu tên)	Không hoặc một lần trong mỗi mục lược đồ phân loại	A.1
<i>Classification Scheme Item Relationship – type description</i> (mối quan hệ mục lược đồ phân loại – mô tả kiểu)	Một lần trong mỗi mục lược đồ phân loại quan hệ kiểu mô tả	A.1
<i>Administration Record – item identifier</i> (bản ghi quản trị – thẻ định danh mục)	Một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – registration status</i> (bản ghi quản trị – tình trạng đăng ký)	Một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – administrative status</i> (bản ghi quản trị – tình trạng quản trị)	Một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – creation date</i> (bản ghi quản trị – ngày tạo)	Một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – last change date</i> (bản ghi quản trị – ngày thay đổi cuối cùng)	Không hoặc một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – effective date</i> (bản ghi quản trị – ngày có hiệu lực)	Không hoặc một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – until date</i> (bản ghi quản trị – ngày có hết hiệu lực)	Không hoặc một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – change description</i> (bản ghi quản trị – mô tả thay đổi)	Không hoặc một lần trong mỗi lược đồ phân loại	A.4
<i>Administration Record – administrative note</i> (bản ghi quản trị – chú thích quản trị)	Không hoặc một lần trong mỗi lược đồ phân loại	A.4

TCVN 7789-2 : 2007

<i>Administration Record – explanatory comment</i> (bản ghi quản trị – lời chú giải)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.4
<i>Administration Record – unresolved issue</i> (bản ghi quản trị – ấn phẩm chưa xử lý)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.4
<i>Administration Record – origin</i> (bản ghi quản trị – nguồn gốc)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.4
<i>Reference Document – identifier</i> (tài liệu tham khảo – thẻ định danh)	Một lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Reference Document – type description</i> (tài liệu tham khảo – mô tả kiểu)	Không hoặc một lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Reference Document – language identifier</i> (tài liệu tham khảo – thẻ định danh ngôn ngữ)	Không hoặc nhiều lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Reference Document – title</i> (tài liệu tham khảo – tiêu đề)	Không hoặc một lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Reference Document – organization name</i> (tài liệu tham khảo – tên tổ chức)	Một hoặc nhiều lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Reference Document – organization mail address</i> (tài liệu tham khảo – địa chỉ thư tín của tổ chức)	Không hoặc một lần trong mỗi <i>tài liệu tham khảo</i>	A.3
<i>Submission – organization name</i> (đệ trình – tên tổ chức)	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Submission – organization mail address</i> (đệ trình – địa chỉ thư tín của tổ chức)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Submission – contact</i> (đệ trình – điểm liên lạc)	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Stewardship – organization name</i> (người quản lý - tên tổ chức)	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Stewardship – organization mail address</i> (người quản lý - địa chỉ thư tín của tổ chức)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Stewardship – contact</i> (người quản lý - điểm liên lạc)	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Registration Authority – organization name</i> (tổ chức có thẩm quyền đăng ký – tên tổ chức)	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Registration Authority – organization mail address</i> (tổ chức có thẩm quyền đăng ký – địa chỉ thư tín của tổ chức)	Không hoặc một lần trong mỗi <i>lược đồ phân loại</i>	A.3
<i>Registration Authority – documentation identifier</i>	Một lần trong mỗi <i>lược đồ phân loại</i>	A.3

(tổ chức có thẩm quyền đăng ký – thể định danh tài liệu)

Registration Authority – documentation language identifier Một hoặc nhiều lần trong mỗi lược đồ phân loại A.3

(tổ chức có thẩm quyền đăng ký – thể định danh ngôn ngữ tài liệu)

Registrar – identifier Một hoặc nhiều lần trong mỗi lược đồ phân loại A.3
(Người giữ sổ đăng ký – thể định danh)

Registrar – contact Một hoặc nhiều lần trong mỗi lược đồ phân loại A.3
(Người giữ sổ đăng ký – điểm liên lạc)

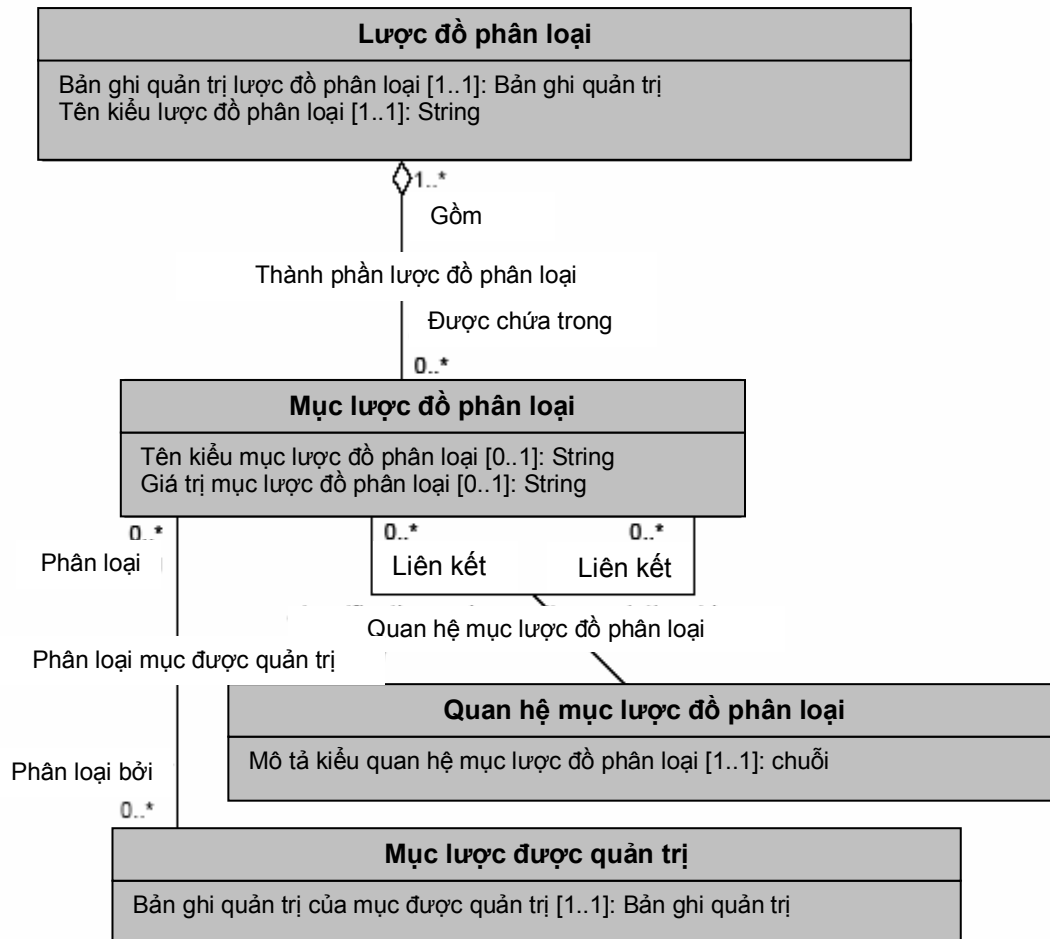
5 Cơ cấu phân loại một lần trong mỗi mục được quản trị

Một lần trong mỗi mục được quản trị trong MDR cần được phân loại bởi lược đồ phân loại bằng cách ghi theo các mối quan hệ sau:

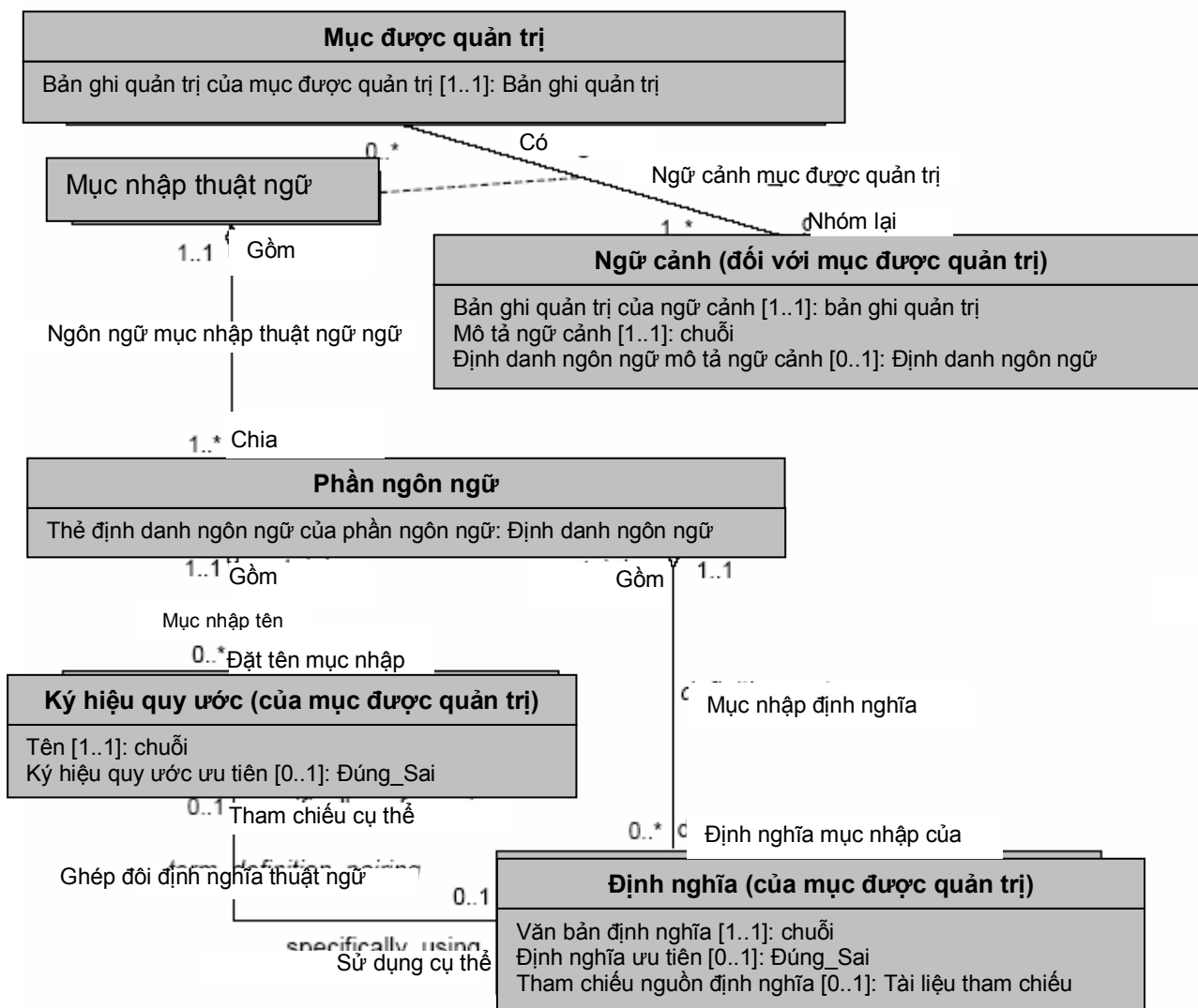
Phân loại mục được quản trị xem hình A.1

Phụ lục A
(Quy định)

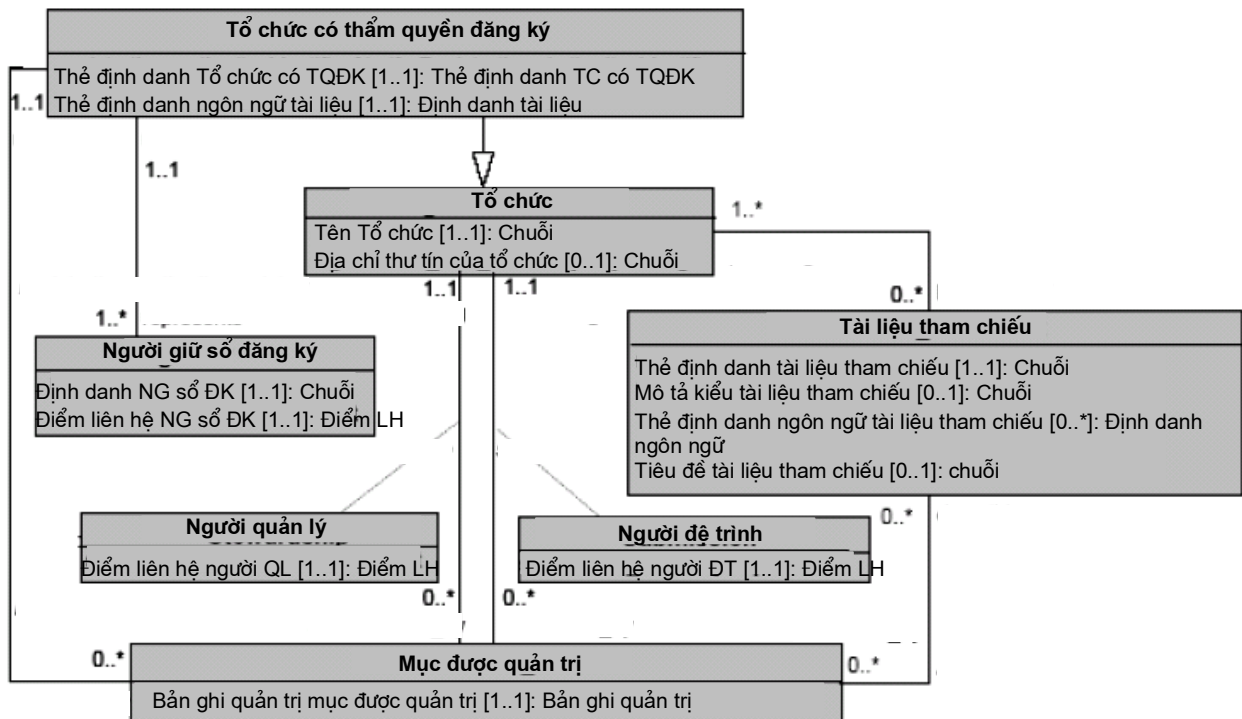
Các lĩnh vực của mô hình dữ liệu MDR bao gồm các thuộc tính của lược đồ phân loại



Hình A.1 - Lĩnh vực phân loại siêu mô hình



Hình A.2 - Phạm vi đặt tên và định nghĩa siêu mô hình



Hình A.3 - Phạm vi quản trị và định danh siêu mô hình

Thẻ định danh tổ chức có thẩm quyền đăng ký		Bản ghi quản trị
Ký hiệu quy ước mã quốc tế [1..1]: Chuỗi Thẻ định danh tổ chức [1..1]: Chuỗi Thẻ định danh bộ phận tổ chức [0..1]: Chuỗi Nguồn OPI [0..1]: Chuỗi		Thẻ định danh mục được đăng ký [1..1]: Thẻ định danh mục Tình trạng đăng ký [1..1]: Chuỗi Tình trạng quản trị [1..1]: Chuỗi Ngày bắt đầu [1..1]: Ngày Ngày thay đổi cuối cùng [0..1]: Ngày Ngày bắt đầu có hiệu lực [0..1]: Ngày Ngày không còn hiệu lực [0..1]: Ngày Mô tả thay đổi [0..1]: Chuỗi Chú thích việc quản trị [0..1]: Chuỗi Ví dụ có bình luận [0..1]: Chuỗi Vấn đề chưa giải quyết [0..1]: Chuỗi Nguồn gốc [0..1]: Chuỗi
Định danh ngôn ngữ		
Thẻ định danh ngôn ngữ [1..1]: Chuỗi Thẻ định danh quốc gia [0..1]: Chuỗi		
Điểm liên lạc		
Tên điểm liên lạc [1..1]: Chuỗi Tiêu đề điểm liên lạc [0..1]: Chuỗi Thông tin liên lạc [1..1]: Chuỗi		
Thẻ định danh mục		
Thẻ định danh tổ chức có thẩm quyền đăng ký mục [1..1]: Thẻ định danh tổ chức có thẩm quyền đăng ký Thẻ định danh dữ liệu [1..1]: Chuỗi Phiên bản [1..1]: Chuỗi		

Hình A.4 - Phạm vi quản trị và định danh siêu mô hình - các lớp được sử dụng như kiểu dữ liệu hỗn hợp

Tài liệu tham khảo

- [1] ISO/IEC Guide 2, Standardization and related activities — General vocabulary (*Tiêu chuẩn hóa và các hoạt động liên quan — từ vựng chung*)
- [2] ISO 1087-1:2000, Terminology work — Vocabulary — Part 1: Theory and application (*thuật ngữ công việc — Từ vựng — Part 1: Lý thuyết và áp dụng*)
- [3] ISO 12620:1999, Computer applications in terminology — Data categories (*Ứng dụng máy tính trong thuật ngữ — Danh mục dữ liệu*)
- [4] ISO 704:2000, Terminology work — Principles and methods (*thuật ngữ công việc — Các nguyên tắc và phương pháp*)
- [5] ISO/IEC 2382-1:1993, Information technology — Vocabulary — Part 1: Fundamental terms (*Công nghệ thông tin — Từ vựng — Phần 1: Các giới hạn quy tắc cơ sở*)
- [6] ISO 2788, Documentation — Guidelines for the establishment and development of monolingual thesauri (*Tài liệu — Hướng dẫn về việc thành lập và phát triển của từ điển đơn ngữ*)
- [7] ISO 5964, Documentation — Guidelines for the establishment and development of multilingual thesauri (*Tài liệu — Hướng dẫn về việc thành lập và phát triển của từ điển đa ngữ*)
-