

TCVN 14171:2024

ĐẶC TẢ TẬP DỮ LIỆU CHO CÔNG DỮ LIỆU MỞ



NGUYỄN TRỌNG KHÁNH

Hà Nội, ngày 25 tháng 11 năm 2025

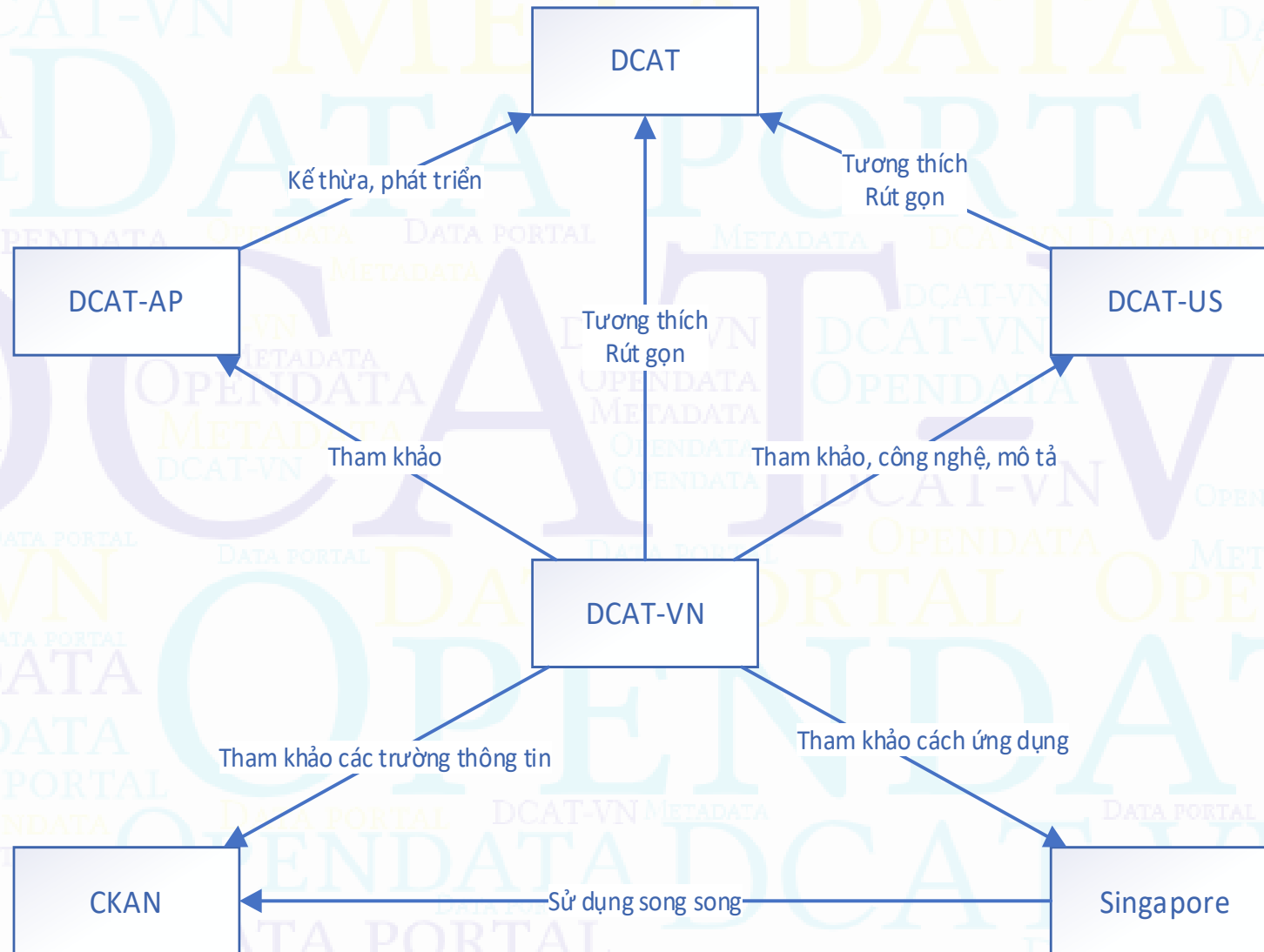
Nội dung

- ❑ Sự cần thiết tiêu chuẩn
- ❑ Mối quan hệ DCAT-VN và các tiêu chuẩn tương tự
- ❑ Nội dung tiêu chuẩn
 - ❖ Vị trí và vai trò
 - ❖ Kết nối giữa các cổng dữ liệu
 - ❖ Cấu trúc đặc tả UML
 - ❖ Mô tả các lớp
 - ❖ Lược đồ
 - ❖ Ví dụ

Sự cần thiết

- ❑ Tập dữ liệu: đơn vị dữ liệu được mở bởi cơ quan, tổ chức
- ❑ Đặc tả tập dữ liệu: thông tin mô tả khái quát về tập dữ liệu như tên, nguồn gốc, cơ quan ban hành, chủ đề, từ khoá...
- ❑ Nếu không có chuẩn:
 - ❖ Cổng dữ liệu nào thì tập dữ liệu chỉ sử dụng cho cổng đó
 - ❖ Không kết nối, liên thông để tra cứu liên cổng
- ❑ Ban hành chuẩn:
 - ❖ Thống nhất các trường, ý nghĩa về tập dữ liệu khi cung cấp
 - ❖ Thuận lợi kết nối, liên thông giữa các cổng dữ liệu (haverst)
 - ❖ Hỗ trợ tìm kiếm trên toàn bộ các cổng có kết nối.

DCAT-VN và các chuẩn tương tự[?]



Nội dung tiêu chuẩn

- ☐ Phạm vi điều chỉnh
- ☐ Tài liệu viện dẫn
- ☐ Mô hình dữ liệu
- ☐ Mô tả chi tiết, ví dụ từng trường
- ☐ Phương pháp kiểm tra (cú pháp, cấu trúc, giá trị)
- ☐ Phụ lục: Mã nguồn lược đồ
- ☐ Phụ lục: Ví dụ mô tả tập dữ liệu XML, JSON
- ☐ Phụ lục: Bảng mã

TCVN

TIÊU CHUẨN QUỐC GIA

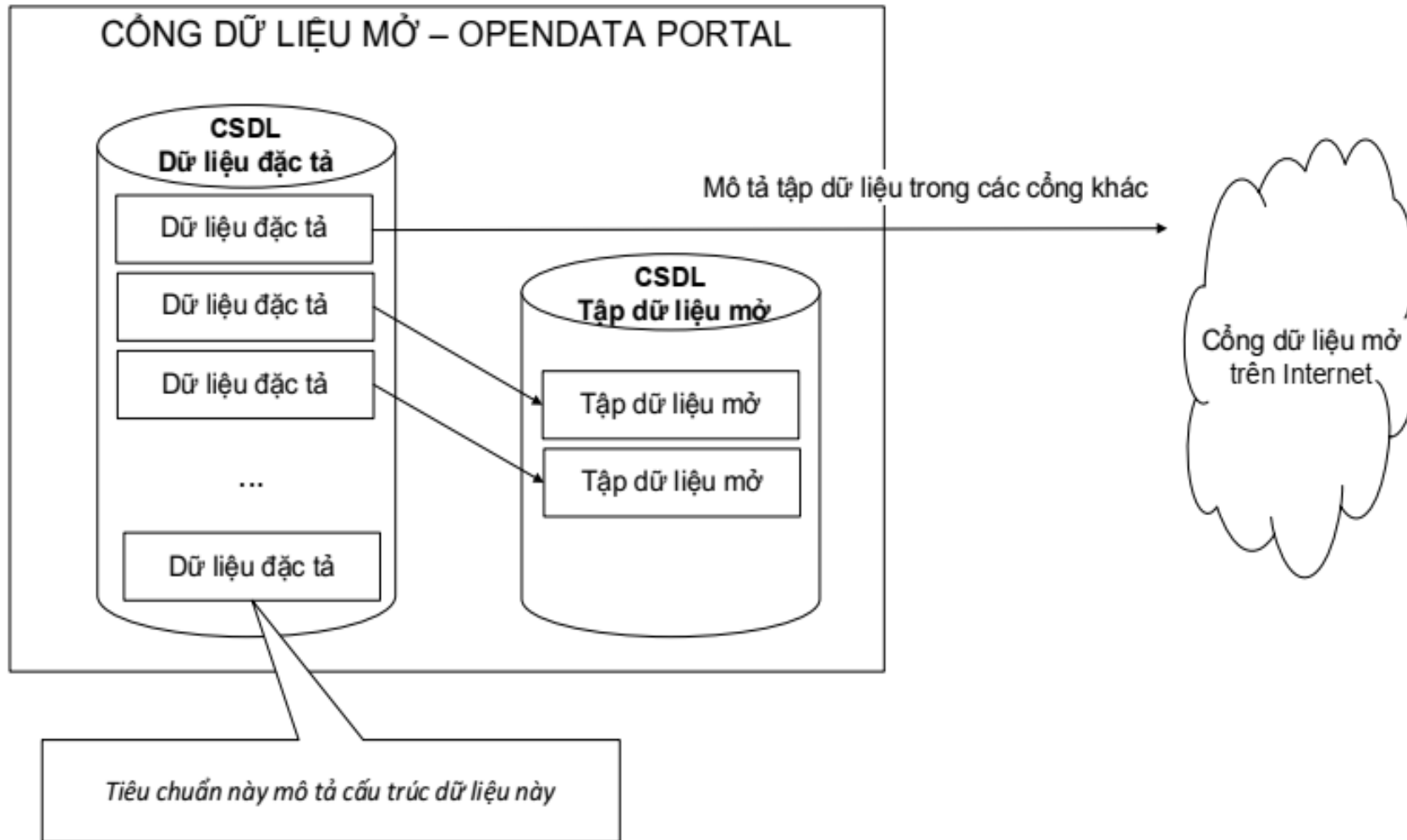
TCVN 14171:2024

Xuất bản lần 1

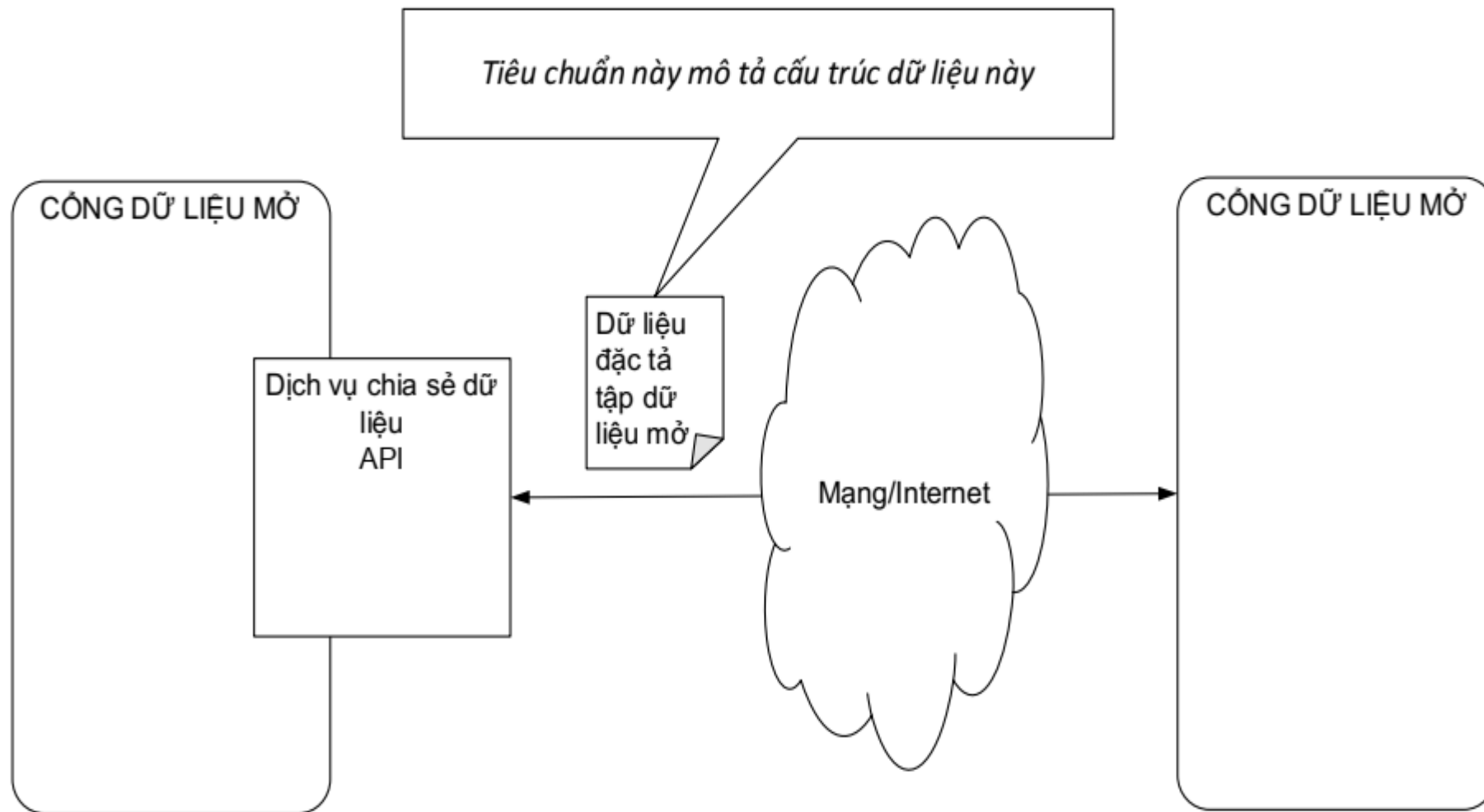
ĐẶC TẢ TẬP DỮ LIỆU MỞ CHO CÔNG DỮ LIỆU (DCAT-VN)

Open dataset specification of data portal (DCAT-VN)

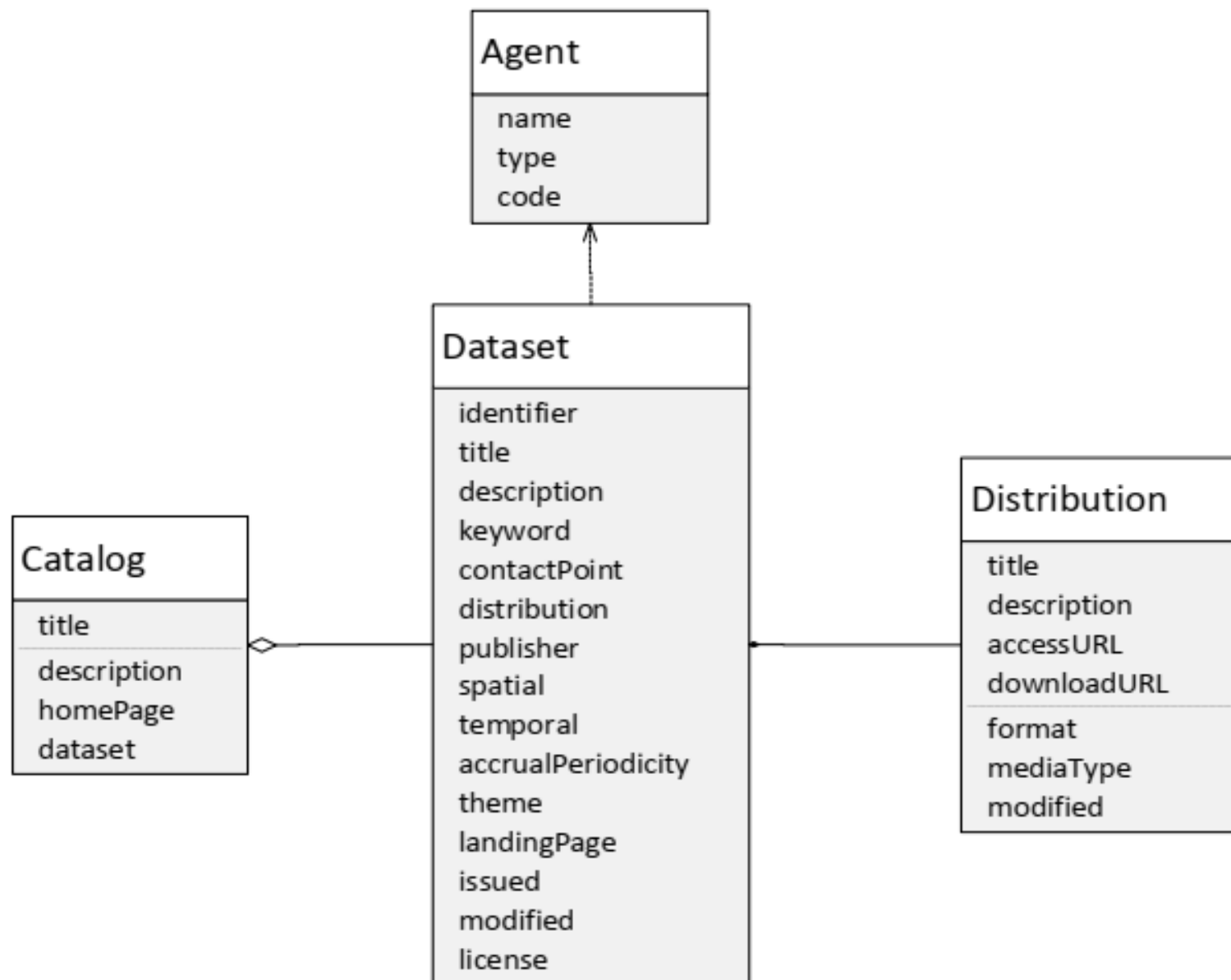
Đặc tả tập dữ liệu mở trong cổng dữ liệu



Kết nối giữa các cổng dữ liệu mở



Cấu trúc đặc tả



Các lớp

- Catalog

- Dataset

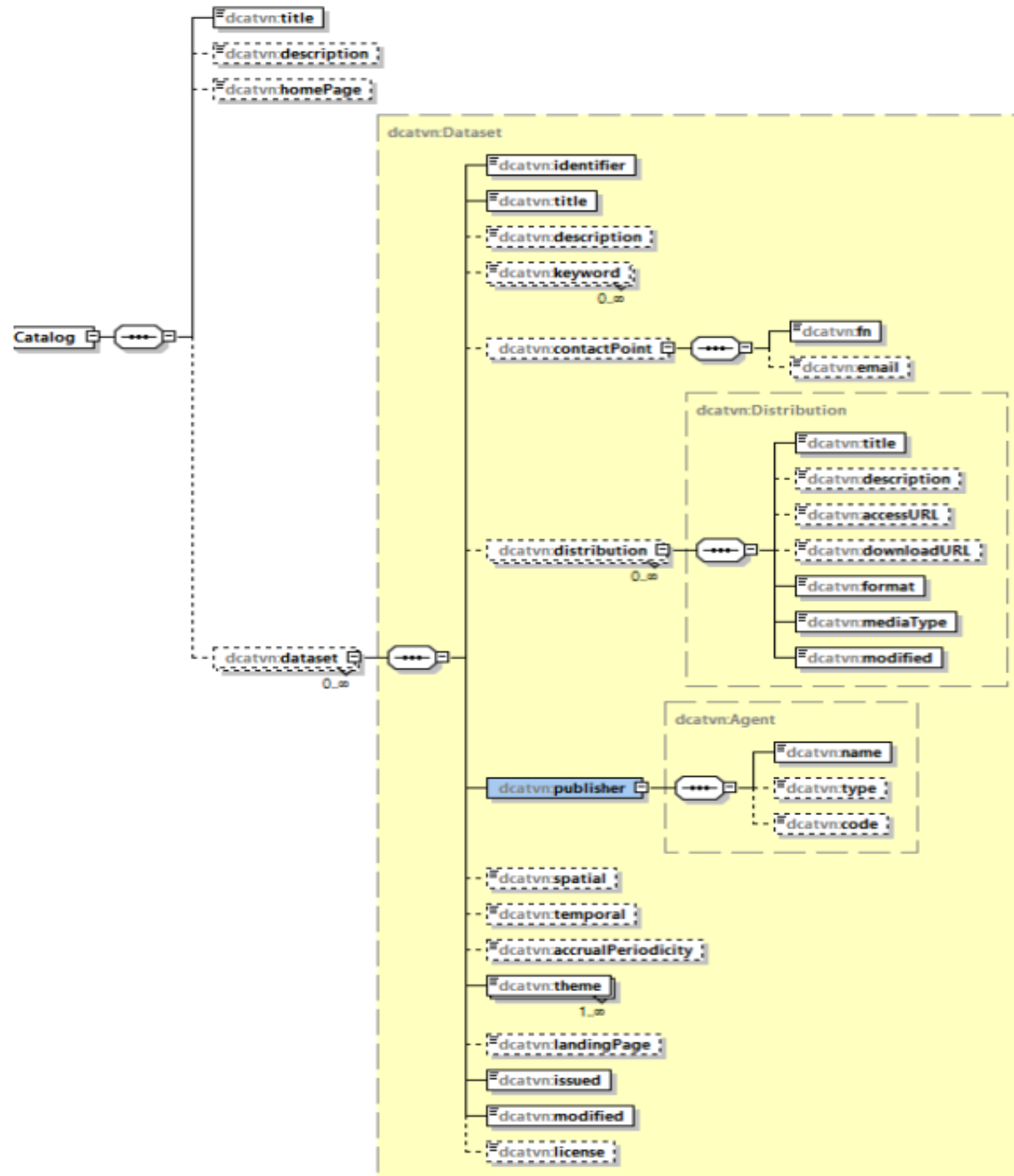
Lớp, trường thông tin		Tên
Catalog		
• title		Tiêu đề
• description		Mô tả
• dataset		Tập dữ liệu
• homePage		Trang chủ của Catalog
Dataset		
▪ title		Tiêu đề
▪ description		Mô tả
▪ keyword		Từ khóa
▪ contactPoint		Liên hệ
◆ fn		Tên liên hệ
◆ email		Địa chỉ email
▪ distribution		Bản phân phối
▪ spatial		Phạm vi địa lý
▪ temporal		Thời gian thu thập
▪ accrualPeriodicity		Tần suất cập nhật
▪ identifier		Định danh tập dữ liệu
▪ theme		Chủ đề
▪ landingPage		Trang chủ của tập dữ liệu được công bố
▪ issued		Thời gian cung cấp
▪ modified		Thời gian cập nhật
▪ licence		Giấy phép
▪ publisher		Cơ quan, tổ chức cung cấp

Các lớp:

- Distribution
- Agent

Distribution	♦ title ♦ description ♦ accessURL ♦ downloadURL ♦ format ♦ mediaType ♦ modified	Tiêu đề Mô tả Địa chỉ truy cập Địa chỉ tải xuống Kiểu dữ liệu Định dạng Thời gian sửa đổi
Agent	♦ name ♦ type ♦ code	Tên cơ quan, tổ chức Loại cơ quan, tổ chức Mã cơ quan, tổ chức

Schema



Mã nguồn lược đồ

A.2 Mã nguồn lược đồ XML Schema

Mã nguồn của lược đồ XML Schema mô tả các lớp dữ liệu, trường dữ liệu của các lớp bằng ngôn ngữ XML. Quy tắc thể hiện qua lược đồ XML.

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://www.w3.org/2001/XMLSchema"
targetNamespace="https://mic.gov.vn/dcat-vn/v1"
xmlns:dcatvn="https://mic.gov.vn/dcat-vn/v1" elementFormDefault="qualified">

  <complexType name="Catalog">
    <sequence>
      <element name="title" type="string" maxOccurs="1"
minOccurs="1"></element>
      <element name="description" type="string" maxOccurs="1"
minOccurs="0"></element>
      <element name="homePage" type="string" maxOccurs="1"
minOccurs="0"></element>
      <element name="dataset" type="dcatvn:Dataset" maxOccurs="unbounded"
minOccurs="0"></element>
    </sequence>
  </complexType>

  <complexType name="Dataset">
    <sequence>
      <element name="identifier" type="string" maxOccurs="1"
minOccurs="1"></element>
      <element name="title" type="string" maxOccurs="1"
minOccurs="1"></element>
      <element name="description" type="string" maxOccurs="1"
minOccurs="0"></element>
      <element name="keyword" type="string" maxOccurs="unbounded"
minOccurs="0"></element>
      <element name="contactPoint" maxOccurs="1" minOccurs="0">
        <complexType>
          <sequence>
            <element name="fn" type="string" minOccurs="1"
maxOccurs="1"></element>
            <element name="email" type="string" minOccurs="0"
maxOccurs="1"></element>
```

Ví dụ (XML)

Ví dụ mã nguồn đặc tả tập dữ liệu mở bằng ngôn ngữ XML

Ví dụ về thông điệp dữ liệu đặc tả tập dữ liệu mở do cơ quan nhà nước cung cấp qua cổng dữ liệu và được trao đổi, chia sẻ giữa các cổng dữ liệu qua dịch vụ chia sẻ dữ liệu. Thông điệp dữ liệu mở được mô tả bằng ngôn ngữ XML.

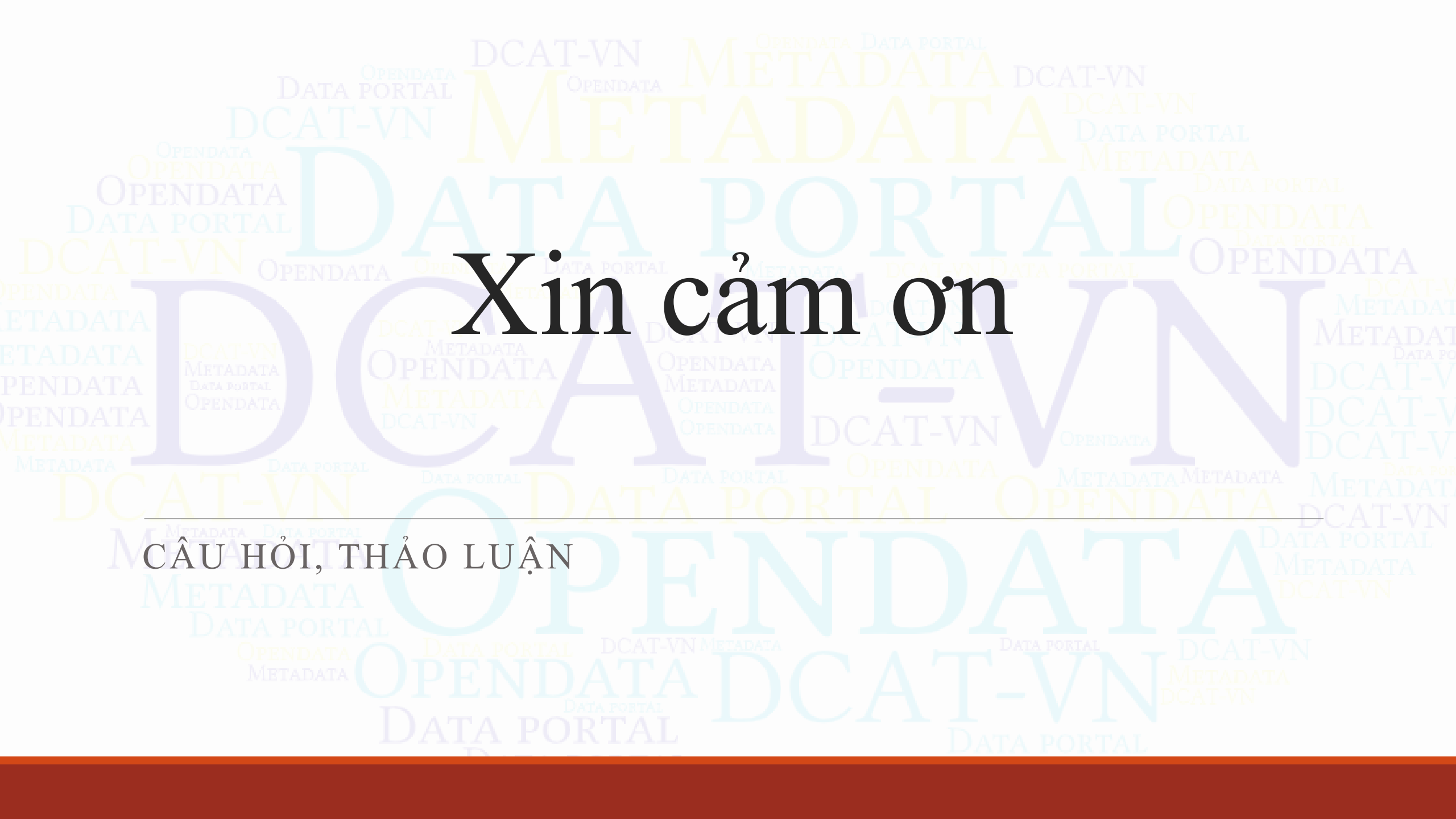
```
<?xml version="1.0" encoding="UTF-8"?>
<Catalog xmlns="https://mic.gov.vn/dcat-vn/v1"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="https://mic.gov.vn/dcat-vn/v1 OpendataMetadata.xsd ">
  <title>Dữ liệu mở thành phố Đà Nẵng 2022</title>
  <description>Các tập dữ liệu mở thành phố Đà Nẵng 2022</description>
  <homePage>https://congdlieu.vn</homePage>
  <dataset>
    <identifier>2ac2b17e-53b4-11ed-bdc3-0242ac120002</identifier>
    <title>Thông tin về tuyến xe buýt hoạt động tại Đà Nẵng</title>
    <description>Thông tin về tuyến xe buýt hoạt động tại Đà Nẵng</description>
    <keyword>Đà Nẵng</keyword>
    <keyword>Giao thông</keyword>
    <keyword>Xe buýt</keyword>
    <contactPoint>
      <fn>Trung tâm Công nghệ thông tin, Sở TTTT Đà Nẵng</fn>
      <hasEmail>info@congdlieu.vn</hasEmail>
    </contactPoint>
    <distribution>
      <title>Danh_sach_tuyen_xe_buýt.csv</title>
      <description>Danh sách các tuyến xe buýt</description>
      <accessURL>https://congdlieu.vn/dulieuchitiet/818574</accessURL>
      <downloadURL>https://open.data.gov.vn/dataset/danh-sach-cac-buu-di-n-buu-c-c-
da-n-ng/resource/3fda2d91-11c7-4e47-b0ae-cbddec560ac5</downloadURL>
      <format>CSV</format>
      <mediaType>text/csv</mediaType>
      <modified>2022-07-27T12:00:00</modified>
    </distribution>
    <publisher>
      <name>Sở Thông tin và Truyền thông Đà Nẵng</name>
      <type>CQNN</type>
      <code></code>
```

Ví dụ (JSON)

Ví dụ mã nguồn đặc tả tập dữ liệu mở bằng ngôn ngữ JSON

Ví dụ về thông điệp dữ liệu đặc tả tập dữ liệu mở do cơ quan nhà nước cung cấp qua cổng dữ liệu và được trao đổi, chia sẻ giữa các cổng dữ liệu qua dịch vụ chia sẻ dữ liệu. Thông điệp dữ liệu mở được mô tả bằng ngôn ngữ JSON

```
{
  "catalog": {
    "title": "Dữ liệu mở thành phố Đà Nẵng 2022",
    "description": "Các tập dữ liệu mở thành phố Đà Nẵng 2022",
    "homePage": "https://congdulieu.vn",
    "dataset": {
      "identifier": "2ac2b17e-53b4-11ed-bdc3-0242ac120002",
      "title": "Thông tin về tuyến xe buýt hoạt động tại Đà Nẵng",
      "description": "Thông tin về tuyến xe buýt hoạt động tại Đà Nẵng",
      "keyword": [
        "Đà Nẵng",
        "Giao thông",
        "Xe buýt"
      ],
      "contactPoint": {
        "fn": "Trung tâm Công nghệ thông tin, Sở TTTT Đà Nẵng",
        "hasEmail": "info@congdulieu.vn"
      },
      "distribution": {
        "title": "Danh_sach_tuyen_xe_buýt.csv",
        "description": "Danh sách các tuyến xe buýt",
        "accessURL": "https://congdulieu.vn/dulieuchitiet/818574",
        "downloadURL": "https://open.data.gov.vn/dataset/danh-sach-cac-buu-di-n-
        buu-c-c-da-n-ng/resource/3fda2d91-11c7-4e47-b0ae-cbddec560ac5",
        "format": "CSV",
        "mediaType": "text/csv",
        "modified": "2022-07-27T12:00:00"
      },
      "publisher": {
        "name": "Sở Thông tin và Truyền thông Đà Nẵng"
```



Xin cảm ơn

CÂU HỎI, THẢO LUẬN