

**TCVN**

**TIÊU CHUẨN QUỐC GIA**

**TCVN 12200:2018**

Xuất bản lần 1

**CÔNG NGHỆ THÔNG TIN - QUY TRÌNH SỐ HÓA VÀ  
TẠO LẬP DỮ LIỆU ĐẶC TẢ CHO ĐỐI TƯỢNG 2D**

*Information technology - Process and metadata creation for 2D digitization*

**HÀ NỘI - 2018**

## Mục lục

1 Phạm vi áp dụng.....	5
2 Tài liệu viện dẫn .....	5
3 Thuật ngữ, Định nghĩa .....	6
12 Dữ liệu đặc tả trong Quy trình số hóa.....	23
13 Quy trình thiết lập dữ liệu đặc tả cho dữ liệu 2D .....	24
13.1 Ước lượng yêu cầu và phân tích nội dung đối tượng .....	25
13.2 Biên tập các yêu cầu chức năng dữ liệu đặc tả.....	25
13.3 Thiết lập hệ thống dữ liệu đặc tả .....	25
13.4 Xây dựng dịch vụ và đánh giá .....	25
14 Lập danh mục liệt kê dữ liệu đặc tả cho dữ liệu 2D .....	25
15 Dữ liệu đặc tả cho đối tượng 2D cơ bản.....	27
15.1 Điểm (Point) .....	27
15.2 Đường (line) .....	27
15.3 Cung tròn (Arc) .....	27
15.4 Đường tròn (Circle).....	28
15.5 Dạng đĩa (Disk).....	28
16 Dữ liệu đặc tả mô tả đối tượng 2D .....	28
Phụ lục A (Tham khảo) Dữ liệu đặc tả mô tả dạng ảnh.....	30
Phụ lục B (Tham khảo) Khuyến nghị đặt tên dữ liệu đặc tả tập tin .....	42
Phụ lục C (Tham khảo) Kiến nghị kiểm soát chất lượng .....	44
Phụ lục D (Tham khảo) Các kỹ năng yêu cầu đối với nhân viên.....	48
Thư mục tài liệu tham khảo .....	49

## **Lời nói đầu**

TCVN 12200:2018 được xây dựng trên cơ sở tham khảo các tiêu chuẩn quốc tế ISO/TR 23081-1:2018 ISO 13028, ISO 19104, ISO 19115-1, ISO 19115-2 và ISO/IEC 19775-1.

TCVN 12200:2018 do Viện Công nghệ thông tin, Đại học Quốc gia Hà Nội biên soạn, Bộ Thông tin và Truyền thông đề nghị, Tổng cục Tiêu chuẩn Đo lường Chất lượng thẩm định, Bộ Khoa học và Công nghệ công bố.

## Công nghệ thông tin - Quy trình số hóa và tạo lập dữ liệu đặc tả cho đối tượng 2D

*Information technology - Process and metadata creation for 2D digitization*

### 1 Phạm vi áp dụng

Tiêu chuẩn này thiết lập các hướng dẫn và nguyên tắc chung đảm bảo cho Quy trình số hóa dữ liệu 2D diễn ra đúng và hiệu quả nhất nhằm đạt được các yêu cầu cần thiết về việc số hóa dữ liệu 2D.

Tiêu chuẩn này đóng vai trò như một hướng dẫn trong việc xây dựng Quy trình số hóa dữ liệu 2D cho tổ chức, doanh nghiệp, thư viện, v.v để đạt được hiệu quả và sự chuẩn xác.

Tất cả các cơ quan, tổ chức, doanh nghiệp đều có thể sử dụng tiêu chuẩn này để thực hiện Quy trình số hóa cho riêng mình.

Tiêu chuẩn này:

- Hướng dẫn để tạo ra và quản lý các bản ghi kỹ thuật số, các tài liệu gốc, hoặc hồ sơ chưa được số hóa khác, đã được sao chép bằng cách số hóa;
- Hướng dẫn thực hành tốt nhất cho việc số hóa để đảm bảo độ chính xác và độ tin cậy của hồ sơ và cho phép xem xét xử lý các hồ sơ gốc;
- Hướng dẫn thực hành tốt nhất cho việc tiếp cận các hồ sơ số hoá khi được yêu cầu;
- Xác định chiến lược để hỗ trợ cho việc tạo ra các hồ sơ phù hợp với mục đích số hóa để duy trì lâu dài;
- Hướng dẫn thực hành tốt nhất cho việc quản lý các hồ sơ nguồn phi kỹ thuật số sau số hóa;
- Các quy trình của tiêu chuẩn này được xây dựng nhằm đáp ứng các yêu cầu đã được xác định qua đánh giá rủi ro;

Tiêu chuẩn này thiết lập các hướng dẫn và nguyên tắc chung đảm bảo cho Quy trình tạo lập Dữ liệu đặc tả diễn ra đúng tiêu chuẩn trong Quy trình số hóa dữ liệu 2D.

Tiêu chuẩn này như một khung tổng quát cho việc tạo lập các bước và các tiêu chuẩn cần thiết trong quá trình tạo lập Dữ liệu đặc tả.

Tiêu chuẩn này chỉ đề cập đến Dữ liệu đặc tả mô tả dữ liệu cơ bản 2D.

Tất cả các tổ chức, doanh nghiệp đều có thể sử dụng tiêu chuẩn này để áp dụng vào Quy trình tạo lập và quản lý Dữ liệu đặc tả cho riêng mình.

### 2 Tài liệu viện dẫn

Các tài liệu viện dẫn sau là rất cần thiết cho việc áp dụng tiêu chuẩn. Đối với các tài liệu viện dẫn ghi năm công bố thì áp dụng 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.

TCVN 6909:2001 (ISO/IEC 10646-1:2000) Công nghệ thông tin - Bộ mã kí tự tiếng Việt 16-bit.

## **TCVN 12200:2018**

Digitization standard process num A000015, 01-03-2013 Government of Alberta (Quy trình tiêu chuẩn số hóa A000015, phát hành ngày 01-03-2013 của chính phủ Alberta).

### **3 Thuật ngữ, định nghĩa**

Tiêu chuẩn này dùng các thuật ngữ, định nghĩa sau đây.

#### **3.1**

##### **Số hóa (digitization)**

Chuyển đổi các tài liệu sang dạng số để xử lý bằng máy tính.

CHÚ THÍCH: Các ví dụ về số hóa bao gồm quét hoặc hình ảnh, chụp ảnh kỹ thuật số của nguồn phi kỹ thuật số hồ sơ, hoặc chuyển đổi các bản ghi âm giọng nói tương tự với phương tiện truyền thông kỹ thuật số.

#### **3.2**

##### **Dữ liệu 2D (2D data)**

Dữ liệu hai chiều được thể hiện dưới dạng ký hiệu, chữ viết, chữ số, hình ảnh, âm thanh, văn bản hoặc XML,...

#### **3.3**

##### **Đánh chỉ mục (indexing)**

Quá trình thiết lập các điểm tiếp cận nhằm tạo thuận lợi cho việc tìm kiếm dữ liệu và/hoặc thông tin.

#### **3.4**

##### **Kiểu dữ liệu (data type)**

Đặc tả vùng giá trị với các thao tác cho phép trên các giá trị trong vùng này.

#### **3.5**

##### **Tập dữ liệu (dataset)**

Tập hợp có thể định danh của dữ liệu.

#### **3.6**

##### **Dự án số hóa (digitization project)**

Truy vấn, sao chụp các hồ sơ phi kỹ thuật số để nâng cao khả năng tiếp cận và tối đa hóa việc tái sử dụng.

#### **3.7**

##### **Dữ liệu đặc tả (metadata)**

Thông tin mô tả các đặc tính của dữ liệu như nội dung, định dạng, chất lượng, điều kiện và các đặc tính khác nhằm tạo thuận lợi cho quá trình tìm kiếm, truy cập, quản lý và lưu trữ dữ liệu.

#### **3.8**

##### **Dữ liệu không gian (spacial data)**

Dữ liệu mô tả các đối tượng trên bề mặt trái đất, dữ liệu không gian được thể hiện dưới dạng hình học, được biểu diễn dưới ba dạng cơ bản là: điểm, đường, vùng.

#### **3.9**

##### **Dữ liệu phi không gian có cấu trúc (structural non-spacial data)**

Các dữ liệu đã được tổ chức theo một cấu trúc thống nhất, bản thân các cấu trúc này không có hoặc ít có sự biến động theo thời gian... Dữ liệu phi không gian có thể có mối quan hệ trực tiếp với dữ liệu không gian hoặc qua trường khóa.

### 3.10

#### Dữ liệu phi cấu trúc (non-structural data)

Dữ liệu ở dạng tự do và không có cấu trúc được định sẵn, ví dụ như: các tệp tin video, tệp tin ảnh, tệp tin âm thanh, đồ họa.

### 3.11

#### Hệ thống thông tin nghiệp vụ (business information system)

Hệ thống tự động tạo ra hay quản lý dữ liệu và hoạt động của một tổ chức

CHÚ THÍCH: Hệ thống thông tin nghiệp vụ là (thường nhiều hoặc liên quan) ứng dụng với mục đích chính là để giao dịch giữa các đơn vị tổ chức và khách hàng của mình, ví dụ như một hệ thống thương mại điện tử, hệ thống quản lý mối quan hệ khách hàng, hệ thống mục đích xây dựng hoặc cơ sở dữ liệu tùy chỉnh, và tài chính, nguồn nhân lực. Hệ thống thông tin doanh nghiệp thường chứa dữ liệu động mà thường cập nhật liên tục. Một hệ thống quản lý tài liệu điện tử và hồ sơ là một loại hình cụ thể của hệ thống thông tin nghiệp vụ với các chức năng chuyên dụng quản lý hồ sơ của tổ chức và các nguồn thông tin.

### 3.12

#### Quy trình nghiệp vụ số hóa (business-process digitization)

Số hóa hồ sơ và thường xuyên đưa vào hệ thống nghiệp vụ, nơi các hoạt động trong tương lai diễn ra trên hồ sơ kỹ thuật số, chứ không phải trên hồ sơ nguồn phi kỹ thuật số.

CHÚ THÍCH: Mục đích để quản lý các hồ sơ, các phiên bản của các bản ghi trên đó hoạt động nghiệp vụ diễn ra, hoặc các hoạt động nghiệp vụ, là phiên bản mà cần phải được quản lý như hồ sơ chính thức. Trong mọi trường hợp, các tổ chức cần phải phân tích các quá trình nghiệp vụ của họ để xác định và quản lý hồ sơ hoạt động.

### 3.13

#### Hủy (destruction)

Quá trình xóa hoặc hủy dữ liệu, có thể tái sử dụng

### 3.14

#### Hồ sơ tự sinh (born digital records)

Bản ghi được tạo ra ở dạng kỹ thuật số, mà không có nguồn dữ liệu bản cứng tương đương

CHÚ THÍCH: dữ liệu số này khác với dữ liệu:

- được tạo ra từ bản cứng;
- bản cứng mà có thể có nguồn gốc từ một nguồn kỹ thuật số nhưng đã được in ra giấy hay chuyển đổi thành tín hiệu tương tự.

### 3.15

#### Xác định giá trị (disposition)

Loạt quá trình liên quan đến việc thực hiện những quyết định về lưu giữ, tiêu hủy hoặc chuyển giao hồ sơ, được quy định bằng văn bản về thẩm quyền xác định giá trị hồ sơ hoặc các công cụ khác.

### 3.16

#### Các bản ghi nguồn chưa số hóa (non-digital source records)

Phạm vi của các quá trình liên quan đến quyết định thực hiện các hồ sơ lưu giữ, tiêu hủy, hoặc chuyển nhượng được ghi lại trong các cơ quan bố trí hoặc các công cụ khác.

## TCVN 12200:2018

**CHÚ THÍCH:** Một bản ghi nguồn phi kỹ thuật số có thể là một hồ sơ gốc hoặc nó có thể đã được tạo ra bởi sao chép, chuyển đổi.

### 3.17

#### **Thể hiện cơ sở (base representation)**

Thể hiện "các đối tượng địa lý đang di chuyển", bằng cách sử dụng các véc-tơ gốc cục bộ và tung độ cục bộ của một đối tượng hình học tại thời gian chuẩn cho trước.

**CHÚ THÍCH 1:** Một đối tượng hình học cố định có thể chịu sự di chuyển hoặc quay nhưng vẫn giữ nguyên số đồng dư với hình vẽ cơ sở của nó.

**CHÚ THÍCH 2:** Các véc-tơ gốc cục bộ và tung độ cục bộ thiết lập nên một hệ tham chiếu tọa độ kỹ thuật (ISO 19111), còn được gọi là hệ tọa độ khung cục bộ hoặc hệ tọa độ khung Euclidean cục bộ.

### 3.18

#### **Đường biên (boundary)**

Tập thể hiện giới hạn của một thực thể.

**CHÚ THÍCH:** Đường biên thường được sử dụng nhiều nhất trong phạm vi hình học, trong đó tập hợp là một tập hợp các điểm hoặc tập hợp các đối tượng đại diện cho các điểm đó. Ở phạm vi khác, thuật ngữ được sử dụng theo phép ẩn dụ để mô tả sự chuyển tiếp giữa một thực thể và phần còn lại của miền ngôn từ của nó.

### 3.19

#### **Mã (code)**

Sự thể hiện một nhân theo một lược đồ được quy định.

### 3.20

#### **Danh sách mã (codelist)**

Miền giá trị bao gồm một mã đối với mỗi giá trị có thể chấp nhận được.

### 3.21

#### **Không gian mã (codespace)**

Quy tắc hoặc quyền của một mã, tên, thuật ngữ hoặc danh mục phân loại.

**VÍ DỤ:** Các ví dụ về không gian mã gồm từ điển, các phân quyền, danh sách mã, ...

### 3.22

#### **Đường cong hỗn hợp (composite curve)**

Chuỗi các đường cong trong đó mỗi đường cong (trừ đường cong đầu tiên) đều bắt đầu từ điểm cuối cùng (của đường cong trước đó trong chuỗi).

**CHÚ THÍCH 1:** Một đường cong hỗn hợp, như là một tập các vị trí trực tiếp, có tất cả các đặc tính của một đường cong.

### 3.23

#### **Hình đặc hỗn hợp (composite solid)**

Tập các Hình đặc được kết nối bằng cách nối liền Hình đặc này với Hình đặc kia dọc theo bề mặt đường biên chung

**CHÚ THÍCH:** Một hình đặc hỗn hợp là một bộ các vị trí trực tiếp, có tất cả các đặc tính của một hình đặc.

### 3.24

#### **Bề mặt hỗn hợp (composite surface)**

Tập các bề mặt được liên kết với nhau bằng cách nối liền bề mặt này với bề mặt kia dọc theo các đường cong đường biên.

**CHÚ THÍCH:** Một bề mặt hỗn hợp là một bộ các vị trí trực tiếp, có tất cả các đặc tính của một bề mặt.

### 3.25

#### Hợp phần (composition)

Dạng kết tập đòi hỏi một trường hợp cụ thể bộ phận phải được chứa trong ít nhất một hợp phần tại một thời điểm, và đối tượng hỗn hợp chịu trách nhiệm tạo thành hoặc phá hủy các bộ phận.

**CHÚ THÍCH:** Các bộ phận có độ bội không cố định có thể được tạo thành sau kết cấu, nhưng ngay khi được tạo thành thì các bộ phận này tồn tại và mất đi với nó (tức là chúng chia sẻ chu trình sống). Các bộ phận như vậy còn có thể bị loại bỏ một cách rõ ràng trước khi hợp phần mất đi. Hợp phần có thể là đệ quy. Từ đồng nghĩa: kết tập hỗn hợp.

### 3.26

#### Biên chung (coboundary)

Tập các nguyên hàm hình học tô-pô về chiều kích hình học tô-pô cao hơn tương ứng với một đối tượng hình học tô-pô nào đó, sao cho đối tượng hình học tô-pô này nằm trong mỗi đường biên của chúng.

**CHÚ THÍCH:** Nếu một nút nằm trên đường biên của một cạnh biên, mà cạnh biên lại nằm trên biên chung của nút đó thì tham số có hướng bất kỳ kết hợp với một trong số các hệ thức này cũng sẽ kết hợp với hệ thức khác. Vì thế nếu nút là nút cuối cùng của cạnh biên [được xác định là đầu của cạnh biên có hướng theo chiều dương], thì hướng theo chiều dương của nút [được gọi là nút có hướng theo chiều dương] sẽ có cạnh biên trên biên chung của nó.

### 3.27

#### Nút kết nối (connect node)

Nút bắt đầu hoặc kết thúc một hoặc nhiều cạnh biên.

### 3.28

#### Tọa độ (coordinate)

Một trong một chuỗi  $n$  số xác định vị trí của một điểm trong không gian  $n$  chiều.

**CHÚ THÍCH:** Trong một hệ tham chiếu tọa độ, các tọa độ được hạn định bằng các đơn vị.

### 3.29

#### Chuyển đổi tọa độ (coordinate conversion)

Thao tác về tọa độ trong đó cả hai hệ tham chiếu tọa độ được dựa trên các mốc tính toán tương tự.

**VÍ DỤ:** Chuyển đổi từ một hệ tham chiếu tọa độ elipxoit dựa trên mốc tính toán WGS84 sang hệ tham chiếu tọa độ Cartesian cũng dựa trên mốc tính toán WGS84, hoặc thay đổi các đơn vị như từ radian sang độ hoặc từ feet sang mét.

**CHÚ THÍCH:** Chuyển đổi tọa độ sử dụng các tham số có giá trị quy định mà không được xác định bằng thực nghiệm.

### 3.30

#### Chiều kích thước tọa độ (coordinate dimension)

Số các phép đo hoặc các trục cần để mô tả một vị trí trong một hệ tọa độ.

### 3.31

#### Thao tác về tọa độ (coordinate operation)

Sự thay đổi các tọa độ, dựa trên một mối quan hệ 1 đổi 1, từ một hệ tham chiếu tọa độ này sang một hệ tham chiếu tọa độ khác.

**CHÚ THÍCH:** Kiểu cơ sở của phép biến đổi tọa độ và sự chuyển đổi tọa độ.

### 3.32



## **TCVN 12200:2018**

### **Điểm (point)**

Nguyên gốc hình học 0- chiều, biểu thị một vị trí.

CHÚ THÍCH: Đường biên của một điểm là tập rỗng.

### **3.33**

#### **Hình bao phủ điểm (Hình bao phủ điểm)**

Hình bao phủ có một miền bao gồm các điểm.

### **3.34**

#### **Đa giác (polygon)**

Bề mặt phẳng được xác định bởi 1 đường biên bên ngoài và 0 hoặc nhiều đường biên phần bên trong.

### **3.35**

#### **Bao phủ đa giác (polygon coverage)**

Hình bao phủ có một miền gồm nhiều đa giác.

### **3.36**

#### **Đa hình (polymorphism)**

Đặc tính của việc có thể gán một ý nghĩa khác hoặc sử dụng một cái gì đó trong các bối cảnh khác nhau một cách đặc biệt, để cho phép một thực thể như là một biến, một chức năng, hoặc một đối tượng để có nhiều hơn một hình dạng

CHÚ THÍCH: Có một vài loại khác về đa hình.

### **3.37**

#### **Chất lượng (quality)**

Toàn bộ các đặc tính của một sản phẩm liên quan tới khả năng nó thỏa mãn được các nhu cầu đã đưa ra và được biết đến.

### **3.38**

#### **Lược đồ chất lượng (quality schema)**

Lược đồ khái niệm xác định các khía cạnh về chất lượng của dữ liệu địa lí.

### **3.39**

#### **Lưới tọa độ (raster)**

Khuôn mẫu hình chữ nhật thông thường của việc tạo thành các dòng quét song song, hoặc tương ứng với màn hình hiển thị trên một đèn điện tử có tia âm cực

CHÚ THÍCH: Lưới tọa độ là một kiểu của lưới.

### **3.40**

#### **Véc-tơ (vector)**

Số lượng có hướng cũng như vớ độ lớn

CHÚ THÍCH: Một đoạn thẳng có hướng biểu thị một véc-tơ nếu độ dài và hướng của đoạn thẳng bằng với cường độ và hướng của véc-tơ. Thuật ngữ dữ liệu véc-tơ đề cập đến dữ liệu trình bày cấu hình không gian của đối tượng địa lý như một tập các đoạn thẳng có hướng.

### **3.41**

#### **Hình học véc- tơ (vector geometry)**

Biểu thị hình học thông qua sử dụng nguyên gốc hình học có cấu trúc.

**3.42****Hệ tham chiếu tọa độ dọc (vertical coordinate reference system)**

Hệ tham chiếu tọa độ một chiều dựa trên một mốc tính toán dọc.

**3.43****Miêu tả hình vẽ (picture portrayal)**

Thể hiện về dữ liệu ảnh dưới dạng các hệ tọa độ màu sắc -không gian phù hợp và kết hợp chặt chẽ giữa các đặc điểm của một thiết bị đầu ra thực hoặc ảo được quy định và việc xem xét hình ảnh.

CHÚ THÍCH: Các bản miêu tả hình vẽ phục vụ cho việc hiển thị trực quan bất kể dù là dưới dạng bản cứng hay bản mềm.

**3.44****Điểm ảnh (pixel)**

Phần tử nhỏ nhất của ảnh kỹ thuật số mà ở đó các thuộc tính được gán vào.

CHÚ THÍCH 1: Thuật ngữ này xuất phát như sự rút gọn của "phần tử hình vẽ".

CHÚ THÍCH 2: Có liên quan đến khái niệm về ô lưới.

**3.45****Nút (node)**

Nguyên hàm hình học tô-pô 0- chiều

CHÚ THÍCH: Đường biên của một nút là tập rỗng.

**3.46****Phần Dữ liệu đặc tả (metadata section)**

Tập con của Dữ liệu đặc tả gồm một tập hợp các thực thể siêu liên quan và các phần tử Dữ liệu đặc tả.

CHÚ THÍCH: Tương đương với một gói trong thuật ngữ UML.

**3.47****Lược đồ Dữ liệu đặc tả (metadata schema)**

Lược đồ khái niệm mô tả Dữ liệu đặc tả.

CHÚ THÍCH: ISO 19115 xác lập một tiêu chuẩn cho lược đồ dữ liệu đặc tả.

**3.48****Thực thể Dữ liệu đặc tả (metadata entity)**

Tập các phần tử Dữ liệu đặc tả miêu tả các khía cạnh của cùng dữ liệu

CHÚ THÍCH 1: Có thể gồm một hoặc nhiều thực thể Dữ liệu đặc tả.

CHÚ THÍCH 2: Tương đương với một lớp trong thuật ngữ UML.

**3.49****Phần tử Dữ liệu đặc tả (metadata element)**

Đơn vị rời rạc của Dữ liệu đặc tả.

CHÚ THÍCH 1: Các phần tử Dữ liệu đặc tả là duy nhất trong một thực thể Dữ liệu đặc tả

CHÚ THÍCH 2: Tương đương với một thuộc tính trong thuật ngữ UML.

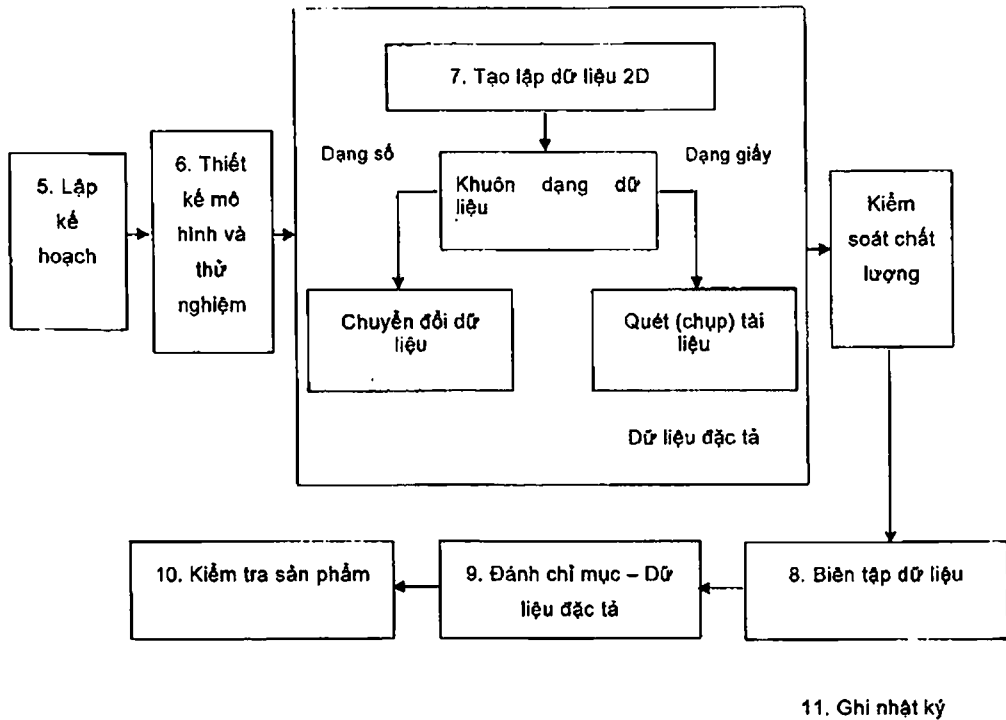
**3.50****Nút giao (junction)**

Nút hình học tô-pô đơn trong một mạng với tập hợp kèm theo là các vòng quay, các nối kết đến và đi

CHÚ THÍCH: Nút giao là một biệt hiệu của nút.

#### 4 Quy trình số hóa dữ liệu 2D

Quy trình số hóa tài liệu trên Hình 1 gồm 8 bước (từ bước 5 đến bước 11), trong đó các bước tương ứng được mô tả cụ thể trong các phần tiếp theo.



Hình 1 - Quy trình số hóa 2D

#### 5 Lập kế hoạch

Tất cả quá trình số hóa phải được lên kế hoạch, xác định phạm vi và ghi lại tài liệu. Các tài liệu dự án bao gồm:

- phạm vi định nghĩa: xác định rõ ràng nghiệp vụ, mục tiêu, quy mô, kích thước và những hạn chế của dự án;
- tuyên bố về mục đích và dự kiến sử dụng của các hồ sơ kỹ thuật số, minh họa nếu cần thiết với các ví dụ;
- tuyên bố về lợi ích, rủi ro: xác định rõ ràng về những lợi ích, rủi ro dự kiến từ việc số hóa;
- báo cáo kết quả nhu cầu và tác động của người dùng: ví dụ, làm thế nào các hồ sơ số hoá sẽ được sử dụng truy cập và ảnh hưởng thế nào người sử dụng;
- tuyên bố về tiêu chuẩn kỹ thuật áp dụng: bao gồm định dạng, nén và Dữ liệu đặc tả;
- thiết bị và các nguồn lực để hỗ trợ cho việc số hóa;

- h) quy trình lập kế hoạch, kiểm soát và thực hiện việc số hóa, bao gồm cả những người thực hiện trước, trong và sau khi số hóa;
- i) quy trình kiểm soát chất lượng;
- k) chiến lược cho việc tích hợp các hình ảnh số hóa vào quy trình làm việc để hỗ trợ các hoạt động nghiệp vụ đang diễn ra;

chiến lược cho quản lý đối với các hồ sơ được số hóa và hồ sơ nguồn phi kỹ thuật số phải luôn sẵn sàng khi được yêu cầu.

### 5.1 Rà soát, phân loại các thông tin dữ liệu

#### a) Mục đích

Rà soát, phân loại và đánh giá chi tiết các thông tin dữ liệu phục vụ xây dựng dữ liệu 2D phù hợp với yêu cầu

#### b) Các bước thực hiện

- Rà soát, đánh giá và phân loại chi tiết dữ liệu đã được chuẩn hóa và chưa được chuẩn hóa
- Chuẩn bị dữ liệu mẫu

#### c) Sản phẩm

- Báo cáo rà soát, phân loại và đánh giá các thông tin dữ liệu
- Bộ dữ liệu mẫu

### 5.2 Phân tích nội dung thông tin dữ liệu

#### a) Mục đích

Phân tích, xác định chi tiết các thông tin dữ liệu phục vụ thiết kế và lập dự toán xây dựng dữ liệu 2D

#### b) Các bước thực hiện

- Xác định danh mục các đối tượng quản lý
- Xác định chi tiết các thông tin cho từng đối tượng
- Xác định chi tiết các quan hệ giữa các đối tượng
- Xác định chi tiết các tài liệu quét
- Xác định khung danh mục dữ liệu 2D, Dữ liệu đặc tả
- Xác định các yếu tố ảnh hưởng đến việc tạo lập dữ liệu 2D

#### c) Sản phẩm

- Danh mục đối tượng quản lý và các thông tin chi tiết
- Danh mục chi tiết các tài liệu quét
- Báo cáo khung danh mục dữ liệu, Dữ liệu đặc tả

### 5.3 Lựa chọn các thông số kỹ thuật

Việc lựa chọn các thông số yêu cầu kỹ thuật và nguyên tắc thực hiện với quy trình số hóa là việc làm cần thiết, trong đó các yêu cầu tiêu chuẩn là điều bắt buộc.

Các tiêu chuẩn khuyến nghị sử dụng gồm:

- định dạng tập tin;
- độ phân giải;

## TCVN 12200:2018

- màu sắc hoặc độ sâu bit;
- nén;
- quản lý màu sắc.

Tiêu chuẩn thông số kỹ thuật đang phát triển nhanh chóng, đặc biệt là trong lĩnh vực kỹ thuật, chức năng của thiết bị để thích ứng tiêu chuẩn này. Việc xem xét các thông số kỹ thuật chủ yếu là để đảm bảo mức độ dễ đọc hoặc khả năng sử dụng hình ảnh số hóa. Những tiêu chí cơ bản sau đây nên được xem xét khi lựa chọn tiêu chuẩn kỹ thuật:

- a) thông số kỹ thuật chất lượng cao nhất cần được đưa vào hỗ trợ quá trình số hóa;
- b) có định dạng mã nguồn mở (có nghĩa là không độc quyền) hoặc các tiêu chuẩn đã công bố, có thể sử dụng, hoặc đã được triển khai rộng rãi;
- c) định dạng không chứa các đối tượng nhúng, hoặc liên kết ra đối tượng bên ngoài phiên bản cụ thể của định dạng;
- d) định dạng được hỗ trợ bởi phần mềm ứng dụng và nhiều hệ điều hành;
- e) định dạng có thể dễ dàng đọc bởi thành phần mở rộng nếu phần mềm cụ thể không có sẵn cho người sử dụng;
- f) hỗ trợ kỹ thuật đầy đủ để cho phép thực hiện bảo trì và khả năng di chuyển khi cần thiết;
- g) các bản sao nên được tạo ra với các tiêu chuẩn kỹ thuật cao nhất được hỗ trợ;
- h) các bản sao chủ nên được lưu giữ lại bất khả xâm phạm, lưu trữ an toàn;
- i) các bản sao phái sinh có thể trở có định dạng thuận tiện nhất cho mục đích kinh doanh (ví dụ như hình thu nhỏ để phân phối trên Internet, vv).

### 5.4 Thiết bị và phần mềm

Thiết bị và phần mềm cần phải được lựa chọn phù hợp với các nguyên tắc, yêu cầu thực hiện số hóa. Chất lượng của các thiết bị và phần mềm ảnh hưởng đến khả năng gắn kết các tiêu chuẩn kỹ thuật phù hợp với yêu cầu, để đảm bảo tuổi thọ của hình ảnh kỹ thuật số. Trong đó tính đến cả trường hợp các hồ sơ gốc sẽ bị hủy, phải đảm bảo khả năng duy trì lâu dài của những hình ảnh số hóa.

### 5.5 Ứng dụng các kỹ thuật nâng cao

Việc sử dụng công nghệ kỹ thuật cao cho hình ảnh số hóa để tạo ra một sự tương đồng chính xác hơn giữa dữ liệu số hóa và tài liệu gốc.

#### 5.5.1 Nâng cao chất lượng hình ảnh

Trong quá trình số hóa, việc sử dụng kỹ thuật để nâng cao chất lượng hình ảnh, tạo có một sự tương đồng chính xác hơn với dữ liệu nguồn nên được dự liệu. Công nghệ này bao gồm "làm sắc nét" và / hoặc "cắt" những điểm sáng hay bóng, "làm mờ" để loại bỏ các vết xước, xóa nhiễu.

#### 5.5.2 Quản lý chú thích

Đó là nơi mà phần mềm được sử dụng để quản lý ảnh kỹ thuật số sau khi quét để bổ sung chú thích cho hình ảnh, như là làm nổi bật, đóng tem, bổ sung các ghi chú, chú thích, những việc này cũng cần được quản lý cho tốt để không làm thay đổi hình ảnh thực tế.

### 5.5.3 Chất lượng hình ảnh

Hình ảnh kỹ thuật số phải đạt chất lượng theo yêu cầu. Điều này có thể được xử lý nhờ các thông số kỹ thuật của thiết bị. Ví dụ, nếu chất lượng màu sắc của bản đồ là quan trọng thì thiết bị để tạo ra hình ảnh cần phải hỗ trợ và phân tích xử lý được việc này.

### 5.5.4 Phương tiện lưu trữ

Tùy thuộc vào yêu cầu của người sử dụng như thời gian lưu trữ hình ảnh, tốc độ truy vấn ... thì cần sử dụng các phương tiện lưu trữ cho phù hợp.

## 6 Thiết kế mô hình và kiểm thử

### 6.1 Thiết kế mô hình

#### a) Mục đích

Thiết kế mô hình danh mục dữ liệu, Dữ liệu đặc tả dựa trên kết quả rà soát, phân tích.

#### b) Các bước thực hiện

- Thiết kế mô hình danh mục dữ liệu, Dữ liệu đặc tả.
- Nhập dữ liệu mẫu

#### c) Sản phẩm

- Mô hình danh mục dữ liệu, Dữ liệu đặc tả dưới dạng XML.
- Báo cáo thuyết minh mô hình danh mục dữ liệu, Dữ liệu đặc tả

### 6.2 Tạo lập dữ liệu cho danh mục dữ liệu và Dữ liệu đặc tả

#### a) Mục đích

Tạo lập nội dung dữ liệu cho danh mục dữ liệu, Dữ liệu đặc tả dựa trên kết quả rà soát, phân tích và thiết kế.

#### b) Các bước thực hiện

- Tạo lập nội dung cho danh mục dữ liệu
- Tạo lập nội dung cho Dữ liệu đặc tả

#### a) Sản phẩm

- Danh mục dữ liệu 2D, Dữ liệu đặc tả đã nhập đủ nội dung.
- Báo cáo kết quả thực hiện

### 6.3 Lựa chọn phương pháp số hóa

Việc lựa chọn phương pháp số hóa là cần thiết tùy vào quy mô, nhu cầu của mỗi dự án mà có những lựa chọn cho phù hợp. Điều này không những làm giảm chi phí mà còn tối ưu được quy trình cũng như chất lượng và yêu cầu của dự án.

Với phương pháp tiếp cận nào để số hóa thì cũng cần áp dụng các mục sau:

- phương pháp lựa chọn số hóa nên được dự liệu;
- quy trình kiểm soát chất lượng nên được thực hiện;
- phương pháp số hóa nên được thường xuyên xem xét lại cho phù hợp với yêu cầu, phù hợp và hiệu quả chi phí cũng như pháp lý.

## TCVN 12200:2018

### 6.3.1 Tự hoặc thuê dịch vụ số hóa

Tự số hóa sẽ giúp cho tổ chức có được tất cả các thiết bị và chuyên môn cần thiết để số hoá và tích hợp đầu ra số hóa vào hệ thống riêng của họ.

Thuê ngoài là việc ký hợp đồng với 1 bên đối tác để thực hiện việc số hóa trên cơ sở đại diện cho tổ chức này.

### 6.3.2 Quy trình số hóa hàng loạt hoặc theo yêu cầu số hóa

Xử lý số hóa hàng loạt là việc tập hợp các nguồn dữ liệu với số lượng đủ lớn trước khi số hóa. Đối lập với phương pháp này là số hóa hồ sơ theo yêu cầu khi cần thiết.

### 6.3.3 Số hóa tập trung hay không tập trung

Số hóa tập trung là xử lý số hóa cho tất cả hồ sơ tại 1 địa điểm.

Số hóa không tập trung là việc sử dụng các điểm số hóa tại các nơi khác nhau trong cơ quan

## 7 Tạo lập dữ liệu 2D

### 7.1 Xử lý tài liệu gốc

Chuẩn bị tài liệu là quá trình mà các tài liệu gốc được kiểm tra và chuẩn bị cho việc số hóa. Hoạt động chuẩn bị và xử lý tài liệu gốc cần phải được thực hiện. Mục đích của việc số hóa là để tạo ra dữ liệu số một cách chuẩn xác nhất với dữ liệu gốc. Vì thế cần phải đánh giá trạng thái của dữ liệu gốc, chuẩn hóa cũng như áp dụng các phương pháp cần thiết để đạt được mục tiêu của dự án.

Những hình ảnh đã được số hóa sẽ được sử dụng như là bản sao cho các tài liệu gốc vì thế cần phải đảm bảo chất lượng tốt nhất có thể.

Các hoạt động chuẩn bị, xử lý tài liệu gốc bao gồm:

- a) đánh giá hiện trạng, chất lượng của tài liệu gốc [ví dụ: giấy chất lượng, nhãn, kẹp bìa, các thuộc tính thông tin nội dung (ví dụ như đồ họa)];
- b) các phương pháp số hóa hồ sơ gốc có kích thước phi tiêu chuẩn hoặc các yêu cầu khác (ví dụ số hóa một hồ sơ từ giấy mỏng, tạo một tài liệu kích thước tiêu chuẩn sử dụng sao chụp (photocopy) lớn hoặc nhỏ; cho các hồ sơ mỏng trong các túi bóng mỏng hoặc sử dụng các thiết bị đặc biệt như là máy quét lướt (overhead scanners)) và kiểm tra chất lượng để đảm bảo không có bất kỳ sự mất mát dữ liệu trong quá trình này;
- c) phương pháp xử lý đối với các hồ sơ nguồn gốc chứa các chú thích viết tay, chú thích bên lề, có tẩy bằng bút xóa hoặc các chất khác, đánh dấu hoặc làm bôi đậm các vùng;
- d) phương pháp phân biệt hồ sơ nguồn và hồ sơ bản sao;
- e) phân loại các loại vật liệu không cần được số hóa vì chúng chỉ được dùng tạm thời hoặc ngắn hạn;
- f) chuẩn bị xử lý vật lý cho số hóa (ví dụ như tháo cẩn thận kẹp, liên kết của các trang, sắp xếp các tài liệu, thiết lập kỹ thuật...);
- g) quy trình chỉ định liên kết giữa các tài liệu liên quan để bộ tài liệu coi như là duy nhất, khi đó hình ảnh được số hóa sẽ thể hiện một cách trung thực nhất cho tài liệu gốc (ví dụ như một tài liệu và ghi chú đính kèm; một tài liệu kèm theo một file đính kèm, tài liệu in trên cả hai mặt của giấy hoặc có xác nhận ở mặt sau);

- h) quy trình chỉ định liên kết giữa các bản ghi nguồn phi kỹ thuật số và các bản sao kỹ thuật số, liên kết như vậy thường sẽ được ghi lại bằng cách sử dụng giao thức định danh, trong một số ứng dụng công nghệ mã vạch có thể được sử dụng để liên kết giấy và phiên bản số hóa;
- i) thủ tục kiểm tra và xác minh rằng tất cả các những yêu cầu, mục tiêu đặt ra với tài liệu gốc đã được đưa vào trong quá trình số hóa;

nguyên tắc điều chỉnh việc sắp xếp hoặc nhóm các hồ sơ gốc cho phù hợp (ví dụ như theo kích thước, màu sắc, ngày đặt hàng, kích thước giấy, định dạng, tài liệu ngang hay dọc, đơn hay 2 mặt).

## 7.2 Chuyển đổi dữ liệu

### a) Mục đích

Chuyển đổi dữ liệu dạng số đã chuẩn hóa vào cơ sở dữ liệu

### b) Các bước thực hiện

- Đối với dữ liệu không gian dạng số chưa được chuẩn hóa thì việc chuẩn hóa dữ liệu được thực hiện theo các quy định của từng chuyên ngành trước khi thực hiện chuyển đổi vào cơ sở dữ liệu
- Đối với dữ liệu phi không gian dạng số chưa chuẩn hóa:
  - + Chuẩn hóa phông chữ theo TCVN 6909:2001 (ISO/IEC 10646-1 : 2000);
  - + Chuẩn hóa dữ liệu phi không gian theo thiết kế mô hình dữ liệu 2D.

### c) Sản phẩm

- Dữ liệu dạng số trước khi chuyển đổi.

## 7.3 Quét (chụp) tài liệu

### - Mục đích

Quét (chụp) các tài liệu để phục vụ đính kèm vào các trường thông tin cho các lớp, bảng dữ liệu của đối tượng quản lý.

### - Các bước thực hiện

- + Quét (chụp) các tài liệu
- + Xử lý và đính kèm tài liệu quét.

### - Sản phẩm

Danh mục các tài liệu quét và đã được đính kèm vào các lớp, bảng dữ liệu của các đối tượng quản lý

Quá trình quét chụp ảnh các tài liệu gốc cần số hóa. Quá trình này cần đảm bảo các yêu cầu tiêu chuẩn kỹ thuật đối với ảnh đầu ra. Quá trình quét được thực hiện lại chỉ khi chất lượng hình ảnh không đạt chất lượng hoặc đảm bảo chất lượng kiểm tra. Nếu những hình ảnh số hóa đòi hỏi phải quét lại thì sau đó hình ảnh lại phải đánh chỉ mục và Dữ liệu đặc tả. Quá trình lại được tiến hành thông qua kiểm soát chất lượng, và đảm bảo chất lượng một lần nữa.

## 8 Biên tập dữ liệu 2D

### a) Mục đích

Biên tập dữ liệu 2D theo quy định

### b) Các bước thực hiện

- Hiệu đính nội dung



**TCVN 12200:2018**

**c) Sản phẩm**

- Dữ liệu 2D đã được biên tập

**9 Đánh chỉ mục – Dữ liệu đặc tả**

**9.1 Đánh chỉ mục**

Đánh chỉ mục được sử dụng để phân loại và cung cấp chỉ mục để tách biệt các mục dễ dàng hơn. Thường thì một chỉ mục được phân loại theo ký tự hoặc số.

Quá trình số hóa bao gồm bốn giai đoạn mà đánh chỉ mục phải được áp dụng. Các giai đoạn này là:

- chụp ảnh (quét);
- hình ảnh chụp lại (tái quét);
- kiểm soát chất lượng;
- truyền tải dữ liệu.

Có hai loại thông tin đánh chỉ mục: Thông tin tiêu sử và thông tin thư mục. Thông tin tiêu sử giao dịch với vòng đời của các tập tin hình ảnh, và liên quan đến bối cảnh của các thuộc tính hình ảnh và tập tin đó phải được giữ lại, đăng nhập và xác nhận trong quá trình số hóa.

Các định nghĩa về nghĩa vụ trong đánh chỉ mục bao gồm:

- bắt buộc – thuộc tính phải có;
- bắt buộc nếu có - thuộc tính phải được cung cấp, nếu phù hợp với bối cảnh công việc và / hoặc các nguồn lực (đối tượng nghiệp vụ);
- đề nghị - nên được sử dụng nếu phù hợp với bối cảnh kinh doanh và / hoặc các nguồn lực (đối tượng kinh doanh) ;
- tùy chọn – tùy thuộc vào yêu cầu mà có lựa chọn cụ thể.

**9.1.1 Đánh chỉ mục tiêu sử**

Các thông tin liên quan đến đánh chỉ mục tiêu sử được mô tả trong Bảng 1.

**Bảng 1 – Đánh chỉ mục tiêu sử**

<b>Nghĩa vụ</b>	<b>Quá trình số hóa</b>	<b>Các thuộc tính đánh chỉ mục số hóa</b>
Bắt buộc	Quét	- Hình ảnh liên quan; - Ngày và thời gian số hóa (Lưu ý: thời gian được khuyến khích, nhưng không bắt buộc); - Số lượng trang được số hóa; - Người vận hành thiết bị số hóa và tên thiết bị; - Thông tin tham khảo chéo về hồ sơ hình ảnh.
Bắt buộc	Quét lại (nếu quá trình này là cần thiết)	- Hình ảnh liên quan; - Ngày và thời gian số hóa (Lưu ý: thời gian được khuyến khích, nhưng không bắt buộc); - Số lượng trang được số hóa; - Người vận hành thiết bị số hóa và tên thiết bị; - Thông tin tham khảo chéo về hồ sơ hình ảnh.
Bắt buộc	Đảm bảo chất lượng (khi hoàn thành)	- Tài liệu tham khảo hàng loạt (bắt buộc cho hàng loạt đầu vào);

		- Người thực hiện đảm bảo chất lượng; - Ngày phê duyệt kiểm tra đảm bảo chất lượng.
Bắt buộc nếu có thể	Chuyển dữ liệu	- Ngày chuyển - Tiêu đề chuyển - Chuyển mô tả - Lý do chuyển - Tiếp nhận chuyển

Một số hoặc tất cả các thông tin này có thể được lấy tự động bởi phần mềm số hóa. Trong trường hợp chuyển giao để đảm bảo lưu trữ, điều quan trọng là phải biết ngày chuyển giao để xác định bất kỳ sự chậm trễ nào đó có hợp lý không. Tổ chức cần thực hiện các thủ tục để ngăn chặn bất kỳ hình thức sửa đổi sau khi hình ảnh đã được lấy một cách phù hợp và lập chỉ mục.

### 9.1.2 Đánh chỉ mục thư mục

Thông tin thư mục liên quan đến nội dung và bối cảnh của hồ sơ. Các quyết định lưu giữ và lấy thông tin phải được thực hiện ở phần đầu của một dự án. Thông tin này phải được giữ lại và sau đó kết hợp với các hình ảnh, tốt nhất là bằng các phương tiện tự động hoặc bằng cách nhập dữ liệu thủ công (xác minh kép yêu cầu, trong đó các sản phẩm được nhập hai lần để đảm bảo độ chính xác) trong quá trình số hóa. Các điều khoản thông tin, thư mục sau đây là những yêu cầu cần được lập chỉ mục được mô tả trong Bảng 2.

Bảng 2 – Đánh chỉ mục thư mục

Nghĩa vụ	Các thuộc tính đánh chỉ mục số hóa
Bắt buộc	- Người tạo - Ngày tạo - Ngày thay đổi - Tiêu đề chính thức - Độ nhạy
Bắt buộc (nếu có thể)	- Tiêu đề phụ - Mô tả - Định dạng - Ngôn ngữ - Hoàn cảnh - Định danh
Đề nghị	- Loại - Chủ đề
Tùy chọn	- Mức độ

### 9.2 Dữ liệu đặc tả

Dữ liệu đặc tả được sử dụng để mô tả các đối tượng, các thông tin cần thiết để lưu trữ.

Tất cả hình ảnh được số hóa nên được chỉ định Dữ liệu đặc tả cho quá trình số hóa tài liệu và hỗ trợ các quy trình nghiệp vụ đang diễn ra. Các tổ chức, cơ quan có thể đưa ra các yêu cầu cụ thể và để tối đa hóa sự kế thừa các giá trị dữ liệu từ các hệ thống và thiết bị hiện có. Quy trình quản lý Dữ liệu đặc

## **TCVN 12200:2018**

tả nên tối đa hóa tự động chụp Dữ liệu đặc tả, giảm thiểu việc xử lý thủ công. Bất kỳ việc sử dụng, áp dụng Dữ liệu đặc tả nên được thực hiện có sự tham khảo tiêu chuẩn ISO 23081-1.

Dữ liệu đặc tả kết hợp với hình ảnh là một thành phần thiết yếu trong việc quản lý và truy vấn các hình ảnh.

Hai loại Dữ liệu đặc tả nên thu nhận:

- Dữ liệu đặc tả cụ thể cho các hình ảnh cụ thể và quá trình xử lý ảnh;
- Dữ liệu đặc tả về hồ sơ công việc đang được giao dịch và đại lý liên quan đến nghiệp vụ. Phần lớn các Dữ liệu đặc tả này có thể được tự động sinh ra bởi các phần mềm và phần cứng được sử dụng để quản lý quá trình số hóa. Cần được giảm thiểu càng nhiều càng tốt việc xử lý thủ công.

Dữ liệu đặc tả có thể được nhúng với các nguồn tài nguyên tại thông tin tiêu đề, hoặc có thể được quản lý trong một hệ thống riêng biệt, hoặc cả hai, nhưng trong cả hai trường hợp đó phải có một mối quan hệ trực tiếp hoặc liên hệ giữa chúng; tức là khi Dữ liệu đặc tả nằm trong một hệ thống riêng biệt, nó có liên kết trực tiếp đến các hồ sơ. Dữ liệu đặc tả cũng có thể được đóng gói trong các định dạng hình ảnh.

### **9.3 Các lưu ý Dữ liệu đặc tả cho dự án số hóa**

Trong trường hợp việc truy cập vào nội dung là chủ yếu thì cần tập trung hơn với việc đánh chỉ mục các điểm thường xuyên được tìm kiếm. Các hình ảnh có thể được quản lý giống như các bản ghi riêng biệt, chứ không chỉ đơn giản là các trường liên kết, đặc biệt là nếu có ý định để người dùng từ bên ngoài truy cập thông qua web. Các tổ chức nên tập trung vào đánh chỉ mục các trường cần thiết để tạo thuận lợi cho việc truy cập tới nội dung ở mức sâu hơn.

## **10 Kiểm tra sản phẩm**

### **a) Mục đích**

Kiểm tra dữ liệu 2D đã được tạo lập đảm bảo tính đầy đủ, chính xác, phù hợp với nội dung đã được phê duyệt

### **b) Các bước thực hiện**

- Kiểm tra mô hình dữ liệu 2D
- Kiểm tra nội dung dữ liệu 2D
- Kiểm tra danh mục dữ liệu, Dữ liệu đặc tả

### **c) Sản phẩm**

- Báo cáo kết quả kiểm tra sản phẩm
- Báo cáo kiểm tra, nghiệm thu chất lượng.

## **11 Ghi nhật ký**

Để đảm bảo cho khả năng để xem xét từng giai đoạn, theo dõi hồ sơ và quy trình thì việc sử dụng lưu vết nên được thực hiện trong suốt quá trình số hóa.

Dữ liệu số hóa và chưa số hóa để đảm bảo chất lượng luôn phải đi kèm với 1 hệ thống quản lý chuẩn mực và đảm bảo được các yêu cầu bắt buộc.

### 11.1 Lưu trữ ảnh

Một khi hình ảnh đã đáp ứng tất cả yêu cầu kiểm tra chất lượng, nó phải được đưa vào lưu trữ ngay lập tức.

Tuỳ theo yêu cầu của dự án và chính sách quản lý hợp pháp hồ sơ, một số quá trình số hóa có thể yêu cầu một bước xác nhận mà hình ảnh được số hóa được chứng nhận như là đầy đủ và chính xác trước khi nó có thể được coi như là bản cuối cùng.

Hình ảnh số hóa có thể được chuyển đến thiết bị lưu trữ trung chuyển trước khi được chuyển đến vị trí lưu trữ cuối cùng, nhưng thậm chí lưu trữ tạm thời cũng phải được an toàn.

### 11.2 Chiến lược quản lý dữ liệu

Bất kỳ giải pháp lưu trữ hay quản lý hệ thống nào cũng phải đảm bảo:

- các dữ liệu đã số hóa phải đảm bảo bảo mật và cố định trong tất cả các phương tiện lưu trữ;
- kiểm soát an ninh và truy cập vào phương tiện lưu trữ, phải có khả năng phát hiện và lưu vết các đăng nhập trái phép;
- thời gian phục hồi trong thiết bị lưu trữ ngoại tuyến phải chấp nhận được đối với các nghiệp vụ đang diễn ra;
- dữ liệu số hóa cần chia sẻ để cho phép thực hiện các quá trình xử lý.

#### 11.2.1 Hệ thống quản lý dài hạn

Hệ thống quản lý dài hạn cần phải được thực hiện cho cả hai hồ sơ số hóa và hồ sơ gốc. Phương tiện lưu trữ và các thủ tục nên được xác định, liệt kê và thực hiện. Hiện nay thì chiến lược lưu trữ dài hạn có khả năng để đảm bảo việc quản lý và duy trì hồ sơ số 1 cách đáng tin cậy để có thể đáp ứng nhu cầu của tổ chức. Tuy nhiên, hồ sơ số hoá có thể chiếm không gian lưu trữ đáng kể tùy thuộc vào tỷ lệ chất lượng, độ phân giải và độ nén sử dụng.

#### 11.2.2 Lưu trữ ngắn hạn

Trong trường hợp hệ thống thông tin nghiệp vụ không đáng tin cậy và chuẩn xác hoặc dữ liệu số hóa đầu ra đang được lưu trữ tạm trước khi chuyển sang một hệ thống khác, chiến lược cho việc lưu trữ ngắn hạn có thể bao gồm:

- một máy chủ lưu trữ hoặc giải pháp lưu trữ hồ sơ số hóa khác;
- lưu hồ sơ để số hóa bằng từ;
- lưu hồ sơ số hoá vào các phương tiện lưu trữ (ghi một lần, đọc nhiều) (ví dụ như một đĩa CD hoặc DVD);
- lưu trữ các hồ sơ số hóa là ổ đĩa cứng gắn ngoài.

### 11.3 Sao lưu

Sao lưu thủ tục nên được xác định, liệt kê và thực hiện. Tất cả hồ sơ số hóa, và Dữ liệu đặc tả liên kết của nó nên được đưa vào trong kế hoạch sao lưu của tổ chức. Thủ tục sao lưu được thiết kế để cung cấp đầy đủ việc sao lưu hàng ngày các bản sao của hồ sơ số hóa để trong trường hợp mất mát hoặc hư hỏng tất cả dữ liệu có thể phục hồi một cách dễ dàng và nhanh chóng.

Chế độ sao lưu phải được đưa ra rõ ràng và các bản sao số hóa phải được duy trì với độ bảo mật cao để đảm bảo tính xác thực của bản ghi khi phục hồi dữ liệu.

## **TCVN 12200:2018**

Tất cả các lỗi hệ thống nên được dự liệu, và sử dụng các bản sao lưu cho mục đích phục hồi nên được kiểm tra xác minh để đảm bảo tính toàn vẹn của các hồ sơ phục hồi.

Các chuyên gia công nghệ thông tin thường sử dụng thuật ngữ "lưu trữ" để mô tả việc sao lưu. Đối với mục đích quản lý hồ sơ, tiến hành sao lưu không phải là một chiến lược lưu trữ hay bảo quản, nó là một công việc liên tục hoặc biện pháp phòng ngừa khắc phục thảm họa.

### **11.4 Quản lý dữ liệu gốc**

Hồ sơ nguồn phi kỹ thuật số nên được quản lý một cách thích hợp cho đến khi được đưa vào bảo quản.

Trường hợp hồ sơ nguồn phi kỹ thuật số được giữ lại vì lý do gì đó chứ không phải để kiểm soát chất lượng, hoặc không được ủy quyền cho việc tiêu hủy, hệ thống điều khiển cần được áp dụng. Những hình ảnh kỹ thuật số và các bản ghi nguồn phi kỹ thuật số nên được liên tục liên kết.

Các hồ sơ gốc nên được tổ chức để tối ưu hóa việc thu hồi và cho phép thực hiện hiệu quả việc quản lý và bố trí quy trình.

Dự án số hoá hiếm khi nào cho phép tiêu hủy của các bản ghi gốc. Sau khi quá trình số hóa, hồ sơ nguồn được trả về nguyên vẹn với vị trí của nó từ trước đó để có thể hoạt động như trước khi số hóa trong trường hợp cần thiết.

#### **11.4.1 Xác định giá trị hồ sơ gốc**

Xác định giá trị các bản ghi nguồn phi kỹ thuật số nên được ủy quyền theo quy định của pháp luật có liên quan và dự liệu.

Đối với hồ sơ gốc khi muốn tiêu hủy thì việc đầu tiên là phải xem xét phù hợp với pháp luật liên quan, ủy quyền pháp lý hoặc yêu cầu sự nhất trí các tổ chức khác.

Mọi quyết định tiêu hủy tài liệu gốc cần được dự liệu và thông tin này có thể được truy cập và yêu cầu thực hiện khi cần thiết.

Việc ủy quyền cho tiêu hủy tài liệu gốc cần được ghi lại trong Dữ liệu đặc tả liên kết những ai cho phép như người đứng đầu tổ chức, cơ quan...

Hồ sơ được số hóa nên được (tái) sản xuất được ở định dạng ban đầu của họ theo yêu cầu, nếu có yêu cầu.

Các tổ chức cần lưu ý các tiêu chí, loại trừ và xác định giá trị sau:

- a) các bản ghi số hóa là bản chính xác và đầy đủ của các nguồn phi kỹ thuật số, ghi lại thay thế nó (bao gồm cả tái tạo màu sắc, nếu có);
- b) các thủ tục kiểm soát chất lượng và chứng nhận phù hợp cho quá trình số hóa phải được xác định, thực hiện và giám sát thường xuyên;
- c) các bản ghi số hóa được định danh riêng biệt và liên kết tới nội dung của nó được tạo ra và sử dụng;
- d) một hệ thống quản lý phù hợp được sử dụng để đảm bảo việc sử dụng liên tục của hồ sơ số hoá;
- e) các bản ghi số hóa đáng tin cậy vào trong hoạt động nghiệp vụ;
- f) các Dữ liệu đặc tả cần thiết cho bản ghi số hóa được tạo ra và duy trì;

- g) chương trình xác định trong tổ chức bao gồm các hệ thống thông tin kinh doanh (s) kết hợp với hồ sơ số hóa;
- h) một chiến lược chuyển đổi và / hoặc bảo quản được xác định, dự liệu và thực hiện cho các hồ sơ kỹ thuật số, bao gồm cả hồ sơ số hóa;
- i) yêu cầu pháp lý hoặc quy định để duy trì các bản ghi với cơ chế không được vi phạm;
- k) không có hành động pháp lý hoặc dự kiến sẽ đe dọa sự tiêu hủy nguồn phi kỹ thuật số;

Nguy cơ thách thức đối với tính xác thực và tính toàn vẹn của hồ sơ số hoá đã được đánh giá và xem xét và có thể chấp nhận đối với tổ chức.

## 12 Dữ liệu đặc tả trong quy trình số hóa

Dữ liệu đặc tả được sử dụng để mô tả các đối tượng, các thông tin cần thiết để lưu trữ.

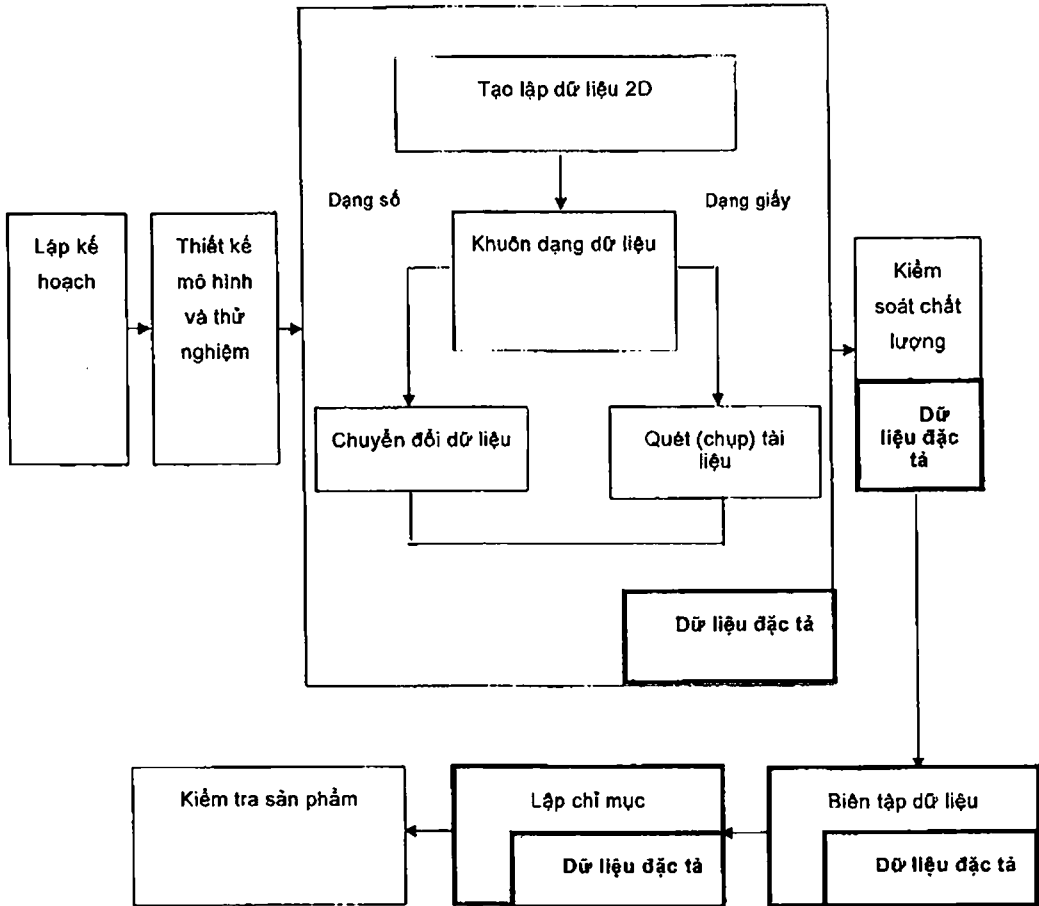
Tất cả hình ảnh được số hóa nên được chỉ định dữ liệu đặc tả cho quá trình số hóa tài liệu và hỗ trợ các quy trình nghiệp vụ đang diễn ra. Các tổ chức, cơ quan có thể đưa ra các yêu cầu cụ thể và để tối đa hóa sự kế thừa các giá trị dữ liệu từ các hệ thống và thiết bị hiện có. Quy trình quản lý dữ liệu đặc tả nên tối đa hóa tự động chụp dữ liệu đặc tả, giảm thiểu việc xử lý thủ công. Bất kỳ việc sử dụng, áp dụng dữ liệu đặc tả nên được thực hiện có sự tham khảo tiêu chuẩn ISO 23081-1: 2018. Dữ liệu đặc tả kết hợp với hình ảnh là một thành phần thiết yếu trong việc quản lý và truy vấn các hình ảnh.

Dữ liệu đặc tả có thể được nhúng với các nguồn tái nguyên tại thông tin tiêu đề, hoặc có thể được quản lý trong một hệ thống riêng biệt, hoặc cả hai, nhưng trong cả hai trường hợp đó phải có một mối quan hệ trực tiếp hoặc liên hệ giữa chúng; tức là khi dữ liệu đặc tả nằm trong một hệ thống riêng biệt, nó cần phải có liên kết trực tiếp đến các hồ sơ. Dữ liệu đặc tả cũng có thể được đóng gói trong các định dạng hình ảnh.

Trong quy trình số hóa bao gồm các giai đoạn mà dữ liệu đặc tả phải được áp dụng. Các giai đoạn này là:

- chụp ảnh (quét);
- hình ảnh chụp lại (tái quét);
- đảm bảo chất lượng;
- biên tập dữ liệu;
- lập chỉ mục.

Hình dưới đây mô tả quy trình số hóa văn bản, trong đó các quá trình bắt buộc áp dụng Dữ liệu đặc tả được đánh dấu bằng nét đậm.



Hình 2 – Dữ liệu đặc tả trong Quy trình số hóa 2D

**13 Quy trình thiết lập Dữ liệu đặc tả cho dữ liệu 2D**

Để xây dựng và triển khai thành công các dự án số hóa và lưu trữ các đối tượng số nói chung và dữ liệu 2D nói riêng, cần thiết lập cơ sở dữ liệu có "các trường và chức năng tích hợp", bằng cách sử dụng việc quản lý tri thức hiệu quả và hiệu lực để tiến hành số hóa, lưu trữ cũng như tìm kiếm và trích xuất tài nguyên một cách nhanh chóng. Trong đó, việc lập kế hoạch Dữ liệu đặc tả cần trọng là cốt yếu để đạt được cách diễn đạt và tìm kiếm hiệu quả dữ liệu 2D. Vì vậy, việc thiết lập Dữ liệu đặc tả là nhiệm vụ cơ bản nhất trong công tác số hóa và lưu trữ kỹ thuật số. Mục đích là để mô tả các nội dung và cấu trúc của các đối tượng số theo một cách thức có cấu trúc và được tiêu chuẩn hóa.

Trước khi xây dựng Dữ liệu đặc tả, cần xem xét và lập kế hoạch cẩn thận, dưới đây là các nguyên tắc trong việc lập kế hoạch đối với Dữ liệu đặc tả:

- Đáp ứng các yêu cầu trong các tiêu chuẩn quốc tế và quốc gia liên quan, bao gồm: khuôn dạng Dữ liệu đặc tả, giao thức truyền thông và các đặc tả tìm kiếm, v.v.;
- Được sửa đổi từ các tiêu chuẩn hiện có, tuân theo các yêu cầu pháp lý của từng quốc gia, và không xây dựng một cách riêng rẽ;
- Dễ sử dụng, bao gồm việc thiết lập dữ liệu, thể hiện, tìm kiếm, dịch lại và tương tác;
- Ứng dụng đa ngôn ngữ;
- Chấp nhận các khuôn dạng Dữ liệu đặc tả khác đối với các lĩnh vực khác và không thiết kế khuôn dạng Dữ liệu đặc tả đơn lẻ có thể áp dụng cho tất cả các lĩnh vực;
- Chấp nhận các khuôn dạng Dữ liệu đặc tả khác, tại cùng thời điểm sử dụng XML để xây dựng cấu trúc Dữ liệu đặc tả để có thể tương thích và tích hợp các khuôn dạng khác nhau, bao gồm cả việc tích hợp các hệ thống hiện có để cho phép tìm kiếm và truy lục toàn bộ các tài nguyên;
- Thiết kế các cơ chế và chức năng quản lý hệ thống đáp ứng với các yêu cầu khác nhau, bao gồm cả việc hạn chế truy cập và bảo vệ quyền sở hữu trí tuệ;
- Khả năng mở rộng Quy trình xử lý, bao gồm cả việc chuyển đổi, lưu trữ, thể hiện và xử lý các khuôn dạng dữ liệu khác nhau cũng như việc thiết lập và chỉnh lý Dữ liệu đặc tả.

Trong các dự án số hóa đối tượng 2D, Quy trình thiết lập Dữ liệu đặc tả cho đối tượng cơ bản được chia thành 4 bước:

#### 13.1 Ước lượng yêu cầu và phân tích nội dung đối tượng

- Phỏng vấn, điều tra các yêu cầu;
- Nghiên cứu các trường hợp sử dụng và các tiêu chuẩn quốc tế/quốc gia về Dữ liệu đặc tả liên quan;
- Phân tích tỉ mỉ các yêu cầu về Dữ liệu đặc tả;
- Xác minh chiến lược Dữ liệu đặc tả và khả năng liên tác giữa các tiêu chuẩn Dữ liệu đặc tả được sử dụng;

#### 13.2 Biên tập các yêu cầu chức năng Dữ liệu đặc tả

- Thiết lập các yêu cầu chức năng Dữ liệu đặc tả;
- Ước lượng và đánh giá hệ thống Dữ liệu đặc tả;

#### 13.3 Thiết lập hệ thống Dữ liệu đặc tả

- Xây dựng các quy phạm thực hành cho việc xây dựng Dữ liệu đặc tả;
- Phát triển và triển khai hệ thống Dữ liệu đặc tả;

#### 13.4 Xây dựng dịch vụ và đánh giá

- Xây dựng các dịch vụ Dữ liệu đặc tả;
- Ước lượng các thao tác Dữ liệu đặc tả;

### 14 Lập danh mục liệt kê Dữ liệu đặc tả cho dữ liệu 2D

Các quy tắc lập danh mục liệt kê cần được thiết lập đầu tiên trước khi lập danh mục liệt kê Dữ liệu đặc tả. Các quy tắc lập danh mục liệt kê nên bao gồm các chi tiết về các phần tử, văn bản, khuôn dạng và các ví dụ minh họa phù hợp.



Bảng sau đây minh họa một danh mục liệt kê các thành phần lõi chung cho đối tượng 2D cơ bản.

**Bảng 3 - Danh mục liệt kê Dữ liệu đặc tả cho dữ liệu 2D**

Phần tử (Yếu tố) Dữ liệu đặc tả		Định nghĩa		
Thông tin định danh	Thông tin trích dẫn	Người khởi tạo	Tên tổ chức/cá nhân tạo dữ liệu 2D cơ bản	
		Ngày tháng công bố	Ngày tháng dữ liệu được công bố	
		Thời gian công bố	Thời gian dữ liệu được công bố	
		Tiêu đề	Tiêu đề đã biết của dữ liệu	
		Phiên bản	Phiên bản của dữ liệu	
		Dạng thể hiện dữ liệu 2D cơ bản	Dạng mà trong đó dữ liệu 2D cơ bản được thể hiện	
		Thông tin về nhóm cùng gốc	Tiêu đề của nhóm gốc mà dữ liệu 2D cơ bản thuộc vào	
		Thông tin công bố	Thông tin công bố được chi tiết	
		Các chi tiết trích dẫn khác	Thông tin khác về trích dẫn	
Thông tin định danh	Thông tin trích dẫn	Trích dẫn công tác rộng hơn		
		Liên kết trực tuyến		
	Mô tả	Mô tả vắn tắt		
		Mục đích		
		Thông tin bổ sung		
	Khoảng thời gian của nội dung đối tượng cơ bản	Thông tin khoảng thời gian	Ngày tháng trong năm	
			Thời gian trong ngày tháng	
		Ngày tháng đơn		
		Thời gian đơn		
		Ngày tháng bắt đầu		
		Thời gian bắt đầu		
		Ngày tháng kết thúc		
		Thời gian kết thúc		
Tham chiếu dòng thời gian				

## 15 Dữ liệu đặc tả cho đối tượng 2D cơ bản

Đối tượng hình học 2D cơ bản chỉ bao gồm các nút hình học do sử dụng hình dạng, thuộc tính hình học và các nút thể hiện bề ngoài được xác định trong thành phần hình dạng. Các nút đối tượng 2D cơ bản có thể được coi là các đối tượng 2 chiều.

### 15.1 Điểm (Point)

#### 15.1.1 Đỉnh (Vertex)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Đỉnh (Vertex)	Tọa độ hai chiều của điểm	(x,y) $X \in (-\infty, \infty); y \in (-\infty, \infty)$

#### 15.1.2 Đa điểm (polyPoint)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Các đỉnh (Vertices)	Bao gồm tập các điểm	[ ] (-∞, ∞)

### 15.2 Đường (line)

#### 15.2.1 Đoạn (Segment)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Điểm đầu (startPoint)	Tọa độ hai chiều của điểm	(x,y) $X \in (-\infty, \infty); y \in (-\infty, \infty)$
Điểm cuối (endPoint)	Tọa độ hai chiều của điểm	(x,y) $X \in (-\infty, \infty); y \in (-\infty, \infty)$

#### 15.2.2 Đa đường (polyLine)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Đoạn (Segment)	Quy định các đỉnh được nối lại với nhau	(x,y) $X \in (-\infty, \infty); y \in (-\infty, \infty)$

### 15.3 Cung tròn (Arc)

#### 15.1.1 Cung mở (Open Arc)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Góc bắt đầu (startAngle)	Giá trị góc bắt đầu quét cung mở theo ngược chiều kim đồng hồ	$-2\pi + 2\pi$
Góc kết thúc (endAngle)	Giá trị góc kết thúc việc quét cung tròn	$-2\pi + 2\pi$
Bán kính (radius)	Bán kính của đường tròn mà cung mở đó là một phần của đường tròn.	$0 + \infty$

#### 15.1.2 Cung đóng (Closure Arc)

Tên dữ liệu đặc tả	Định nghĩa	Miền giá trị
Góc bắt đầu (startAngle)	Giá trị góc bắt đầu quét cung mở theo ngược chiều kim đồng hồ	$-2\pi + 2\pi$
Góc kết thúc (endAngle)	Giá trị góc kết thúc việc quét cung tròn	$-2\pi + 2\pi$

**TCVN 12200:2018**

Bán kính (radius)	Bán kính của đường tròn mà cung đóng đó là một phần của đường tròn.	0 + ∞
Kiểu đóng (ClosureType)	Kiểu đóng cung	{"PIE", "CHORD"}
Đặc (Solid)	Kiểu đặc hoặc rỗng	{"FALSE", "TRUE"}

**15.4 Đường tròn (Circle)**

<b>Tên dữ liệu đặc tả</b>	<b>Định nghĩa</b>	<b>Miền giá trị</b>
Bán kính (radius)	Bán kính của đường tròn.	0 + ∞

**15.5 Dạng đĩa (Disk)**

<b>Tên dữ liệu đặc tả</b>	<b>Định nghĩa</b>	<b>Miền giá trị</b>
Bán kính (innerRadius)	Bán kính trong của đĩa.	0 + ∞
Bán kính (outerRadius)	Bán kính ngoài của đĩa.	0 + ∞
Đặc (Solid)	Kiểu đặc hoặc rỗng	{"FALSE", "TRUE"}

**16 Dữ liệu đặc tả mô tả đối tượng 2D**

Đối tượng hình học được xây dựng dựa trên cơ sở các đối tượng 2D cơ bản. Dữ liệu đặc tả cho đối tượng 2D được hình thành trong quá trình số hóa dữ liệu 2D và tạo lập cơ sở dữ liệu 2D. Dưới đây là các yếu tố dữ liệu đặc tả tối thiểu, áp dụng trong quá trình số hóa cho dữ liệu 2D:

**Bảng 4 – Dữ liệu đặc tả mô tả đối tượng 2D**

Tên dữ liệu đặc tả	Quy định/điều kiện	Lặp lại/Xuất hiện	Kiểu dữ liệu
MD_SpatialRepresentation	Use obligation/condition from referencing object (từ đối tượng tham chiếu)	Use maximum occurrence from referencing object (từ đối tượng tham chiếu)	AggregatedClass (MD_Metadata) "Abstract"
NumberOfDimension	M	1	Integer
axisDimensionProperties	M	N	Sequence
cellGeometry	M	1	Class
transformationParameter Availability	M	1	Boolean
checkPointAvailability	M	1	Boolean

Tên dữ liệu đặc tả	Quy định/điều kiện	Lặp lại/Xuất hiện	Kiểu dữ liệu
checkPointDescription	C/checkPointAvailability = "Yes"	1	characterString
cornerPoint	O	2-4	Sequence
centerPoint	O	1	Class
pointinFixel	M	1	Class
transformationDimension Description	O	1	characterString
transformationDimension Mapping	O	2	characterString
controlPointAvailability	M	1	Boolean
orientationParameter Availability	M	1	Boolean
orientationParameter Description	O	1	characterString
GeoreferenceParameters	M	1	Class
parameterCitation	O	N	Class
MD_vectorSpatialReprentation	Use obligation/condition from referencing object (từ đối tượng tham chiếu)	Use maximum occurrence from referencing object (từ đối tượng tham chiếu)	SpecifiedClass (MD_Spatial Representation)
topologyLevel	O	1	Class
geometricObject	O	N	Class
MD_Dimension	Use obligation/condition from referencing object	Use maximum	Class "Data type"

Tên dữ liệu đặc tả	Quy định/điều kiện	Lặp lại/Xuất hiện	Kiểu dữ liệu
	(từ đối tượng tham chiếu)	occurrence from referencing object (từ đối tượng tham chiếu)	
dimensionName	M	1	Class
dimensionSize	M	1	Integer
resolution	O	1	Class
dimensionTitle	O	1	characterString
dimensionDescription	O	1	characterString
MD_GeometricObject	Use obligation/condition from referencing object (từ đối tượng tham chiếu)	Use maximum occurrence from referencing object (từ đối tượng tham chiếu)	Class "Data type"
geometricObjectType	M	1	Class
geometricObjectCount	O	1	Integer

**Phụ lục A**  
(Tham khảo)

**Dữ liệu đặc tả mô tả dạng ảnh**

Bảng sau đây liệt kê các thành phần bao gồm tất cả các dữ liệu đặc tả mô tả dạng ảnh.

CHÚ THÍCH: Các từ viết tắt sau đây được sử dụng trong cột giao ước của bảng này:

M = Bắt buộc

MA = bắt buộc nếu thích hợp

R = khuyến cáo

O = tùy chọn

**Bảng A.1 – Dữ liệu đặc tả mô tả ảnh**

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lặp lại/xuất hiện
ObjectIdentifier	container	M	Y
objectIdentifierType	string	M	N
objectIdentifierValue	string	M	N
fileSize	positive integer	M	N
FormatDesignation	container	M	N
formatName	string	M	N
formatVersion	string	O	N
FormatRegistry	container	O	N
formatRegistryName	string	O	N
formatRegistryKey	string	O	N
byteOrder	enumerated (list)	MA	N
Compression	container	M	Y
compressionScheme	string	M	N
compressionSchemeLocalList	reference	MA	N
compressionSchemeLocalValue	string	MA	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
compressionRatio	positive integer	O	N
Fixity	container	M	Y
messageDigestAlgorithm	enumerated (list)	M	N
messageDigest	string	M	N
messageDigestOriginator	string	O	N
BasicImageCharacteristics	container	M	N
imageWidth	positive integer	M	N
imageHeight	positive integer	M	N
PhotometricInterpretation	container	M	N
colorSpace	string	M	N
ColorProfile	container	MA	N
IccProfile	container	MA	N
iccProfileName	string	MA	N
iccProfileVersion	string	MA	N
iccProfileURL	string	MA	N
LocalProfile	container	MA	N
localProfileName	string	MA	N
localProfileURL	string	MA	N
embeddedProfile	base64Binary	O	N
YCbCr	container	MA	N
yCbCrSubSampling	Enumerated	MA	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lặp lại/xuất hiện
	(standard)		
yCbCrPositioning	enumerated (standard)	MA	N
yCbCrCoefficients	enumerated (standard)	MA	N
referenceBlackWhite	enumerated (standard)	MA	N
SpecialFormatCharacteristics	container	MA	N
JPEG2000	container	MA	N
CodecCompliance	container	O	N
codec	string	O	N
codecVersion	string	MA	N
codestreamProfile	string	O	N
complianceClass (cClass)	string	O	N
EncodingOptions	container	O	N
tiles	string	O	N
qualityLayers	positive integer	MA	N
resolutionLevels	positive integer	MA	N
MrSID	container	MA	N
zoomLevels	positive integer	MA	N
Djvu	container	MA	N
djvuFormat	enumerated (list)	MA	N



Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
SourceInformation	container	R	N
sourceType	string	R	N
SourceID	container	R	Y
sourceIDType	string	O	N
sourceIDValue	string	O	N
SourceSize	container	O	N
SourceXDimension	container	O	N
sourceXDimensionValue	non-negative real	O	N
sourceXDimensionUnit	enumerated (list)	O	N
SourceYDimension	container	O	N
sourceYDimensionValue	non-negative real	O	N
sourceYDimensionUnit	enumerated (list)	O	N
SourceZDimension	container	O	N
sourceZDimensionValue	non-negative real	O	N
sourceZDimensionUnit	enumerated (list)	O	N
GeneralCaptureInformation	container	MA	N
dateTimeCreated	DateTlme	MA	N
imageProducer	string	R	Y
captureDevice	enumerated (list)	O	N
ScannerCapture	container	MA	N
scannerManufacturer	string	R	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lặp lại/xuất hiện
ScannerModel	container	R	N
scannerModelName	string	R	N
scannerModelNumber	string	R	N
scannerModelSerialNo	string	O	N
maximumOpticalResolution	string	O	N
scannerSensor	enumerated (standard)	R	N
ScanningSystemSoftware	container	R	N
scanningSoftwareName	string	R	N
scanningSoftwareVersionNo	string	R	N
DigitalCameraCapture	container	MA	N
digitalCameraManufacturer	string	R	N
DigitalCameraModel	container	R	N
digitalCameraModelName	string	R	N
digitalCameraModelNumber	string	R	N
digitalCameraModelSerialNo	string	O	N
cameraSensor	enumerated(standard)	R	N
CameraCaptureSettings	container	O	N
ImageData	container	O	N
fNumber	non-negative real	O	N
exposureTime	non-negative real	O	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
exposureProgram	enumerated (standard)	O	N
spectralSensitivity	string	O	Y
isoSpeedRatings	integer	O	N
oECF	rational	O	N
exifVersion	enumerated (standard)	O	N
shutterSpeedValue	rational	O	N
apertureValue	rational	O	N
brightnessValue	rational	O	N
exposureBiasValue	rational	O	N
maxApertureValue	rational	O	N
subjectDistance	non-negative real	O	N
meteringMode	enumerated (list)	O	N
lightSource	enumerated (standard)	O	N
flash	enumerated (standard)	O	N
focalLength	real	O	N
flashEnergy	rational	O	N
backLight	enumerated (standard)	O	N
exposureIndex	non-negative real	O	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lặp lại/xuất hiện
sensingMethod	enumerated (standard)	O	N
cfaPattern	integer	O	N
autoFocus	enumerated (standard)	O	N
PrintAspectRatio	container	O	N
xPrintAspectRatio	non-negative real	O	N
yPrintAspectRatio	non-negative real	O	N
GPSTData	container	MA	N
gpsVersionID	string	MA	N
gpsLatitudeRef	enumerated (standard)	O	N
gpsLatitude	rational	O	N
gpsLongitudeRef	enumerated (standard)	O	N
gpsLongitude	rational		
gpsAltitudeRef	enumerated (standard)	O	N
gpsAltitude	rational	O	N
gpsTimeStamp	string	O	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lặp lại/xuất hiện
gpsSatellites	string	O	N
gpsStatus	enumerated (standard)	O	N
gpsMeasureMode	enumerated (standard)	O	N
gpsDOP	rational	O	N
gpsSpeedRef	enumerated (standard)	O	N
gpsSpeed	rational	O	N
gpsTrackRef	enumerated (standard)	O	N
gpsTrack	rational	O	N
gpsImgDirectionRef	enumerated (standard)	O	N
gpsImgDirection	rational	O	N
gpsMapDatum	string	MA	N
gpsDestLatitudeRef	enumerated (standard)	O	N
gpsDestLatitude	string	O	N
gpsDestLongitudeRef	enumerated (standard)	O	N
gpsDestLongitude	string	O	N
gpsDestBearingRef	enumerated (standard)	O	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
gpsDestBearing	rational	O	N
gpsDestDistanceRef	enumerated (standard)	O	N
gpsDestDistance	rational	O	N
gpsProcessingMethod	string	O	N
gpsAreaInformation	string		
gpsDateStamp	string	O	N
gpsDifferential	enumerated (standard)	O	N
orientation	enumerated (list)	R	N
mcthodology	string	O	N
SpatialMetrics	container	MA	N
samplingFrequencyPlane	enumerated (list)	MA	N
samplingFrequencyUnit	enumerated (standard)	MA	N
xSamplingFrequency	rational	MA	N
ySamplingFrequency	rational	MA	N
ImageColorEncoding	container	M	N
BitsPerSample	container	M	N
bitsPerSampleValue	string	M	N
bitsPerSampleUnit	enumerated (list)	M	N
samplesPerPixel	positive integer	M	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
extraSamples	enumerated (list)	MA	N
Colormap	container	MA	N
colormapReference	reference	MA	N
embeddedColormap	base64Binary	O	N
grayResponseCurve	enumerated (standard)	R	N
grayResponseUnit	enumerated (list)	R	N
WhitePoint	container	O	Y
whitePointXValue	enumerated (list)	O	N
whitePointYValue	enumerated (list)	O	N
PrimaryChromaticities	container	O	Y
primaryChromaticitiesRedX	enumerated (list)	O	N
primaryChromaticitiesRedY	enumerated (list)	O	N
primaryChromaticitiesGreenX	enumerated (list)	O	N
primaryChromaticitiesGreenY	enumerated (list)	O	N
primaryChromaticitiesBlueX	enumerated (list)	O	N
primaryChromaticitiesBlueY	enumerated (list)	O	N
TargetData	container	R	N
targetType	enumerated (list)	R	Y
TargetID	container	R	N
targetManufacturer	string	R	N

Yếu tố dữ liệu đặc tả	Kiểu dữ liệu	Quy định/Điều kiện	Lập lại/xuất hiện
targetName	string	R	N
targetNo	string	R	N
targetMedia	string	R	N
externalTarget	reference	R	Y
performanceData	reference	O	Y
ImageProcessing	container	MA	Y
dateTimeProcessed	DateTime	MA	N
sourceData	string	MA	N
processingAgency	string	R	Y
processingRationale	string	R	N
ProcessingSoftware	container	R	Y
processingSoftwareName	string	R	N
processingSoftwareVersion	string	R	N
processingOperatingSystemName	string	R	N
processingOperatingSystemVersion	string	R	N
processingActions	string	R	Y
Previous Image Metadata	[retains previous data types]	MA	Y



## Phụ lục B

(Tham khảo)

### Khuyến nghị đặt tên Dữ liệu đặc tả tập tin

Phụ lục này là để hướng dẫn việc xác định đặt tên tập tin, đặc biệt cho hình ảnh kỹ thuật số.

#### B.1 Tổng quát

Kế hoạch đặt tên tập tin nên được xây dựng trước khi quét. Sự phát triển của một hệ thống đặt tên tập tin cần tính đến việc định danh do máy hay do con người đánh chỉ số (hoặc cả hai, trong trường hợp này, hình ảnh có thể có nhiều định danh). Tên tập tin phải có ý nghĩa (ví dụ tương quan các tập tin số hóa với các tài liệu gốc), hoặc không được mô tả (như là một chuỗi số tuần tự). Ý nghĩa của tên tập tin chứa Dữ liệu đặc tả là tự tham khảo; tên tập tin không có mô tả mà được kết hợp với Dữ liệu đặc tả được lưu trữ ở nơi khác phục vụ cho việc xác định các tập tin. Nhìn chung, các dự án quy mô nhỏ hơn có thể thiết kế tên tập tin mô tả để tạo điều kiện lựa chọn và truy vấn; dự án quy mô lớn có thể sử dụng tên do máy tính tạo ra và dựa vào một cơ sở dữ liệu tìm kiếm hiện đại và truy vấn Dữ liệu đặc tả liên quan.

Nói chung, tên tập tin nên:

- a) là duy nhất;
- b) có cấu trúc nhất quán;
- c) sử dụng đầu số không để tạo điều kiện sắp xếp theo số thứ tự;
- d) tránh sử dụng các khoảng trắng trong tên tập tin, sử dụng dấu gạch chân như là một thay thế;
- e) tránh sử dụng tên quá phức tạp hay dài để làm người dùng bị lỗi quá trình nhập;
- f) hạn chế độ dài của tên tập tin đến dưới 30 ký tự để tránh các vấn đề tiềm năng với chuyển đổi giữa các
- g) hệ thống khác nhau;
- h) sử dụng các ký tự chữ thường và phần mở rộng tên file;
- i) sử dụng số và / hoặc chữ cái, nhưng không sử dụng các ký tự đặc biệt hoặc khoảng trắng vì có thể gây ra các lỗi với hệ điều hành;

Bản ghi dữ liệu đặc tả được nhúng trong tên tập tin (chẳng hạn như ngày quét, số trang, vv) trong một vị trí khác ngoài các tên tập tin. Điều này cung cấp một mạng lưới an toàn để di chuyển các tập tin giữa các hệ thống trong tương lai, trong trường hợp họ phải được đổi tên. Đặc biệt, thông tin trình tự và các bộ phận kết cấu chính của các đối tượng đa phần phải được ghi rõ ràng trong các dữ liệu đặc tả cấu trúc và không chỉ gắn vào tên tập tin.

## B.2 Cấu trúc thư mục

Các tập tin có thể sẽ được tổ chức trong hệ thống thư mục, tập tin đó sẽ liên kết đến Dữ liệu đặc tả được lưu trữ ở đâu đó trong cơ sở dữ liệu. Các tập tin chính có thể được lưu trữ một cách riêng biệt với các tập tin phái sinh. Thư mục có thể được tổ chức độc lập cho các tập tin hình ảnh, chẳng hạn như các thư mục được sắp xếp theo ngày tháng hoặc cấu trúc phân loại.

Họ cũng có thể nhân rộng quản lý vật lý hay lô-gic cho các hồ sơ nguồn phi kỹ thuật số được quét.

Các tập tin cũng có thể tự được quản lý chỉ bởi cấu trúc thư mục và thư mục trước khi đặt tên có ý nghĩa. Cách tiếp cận này thường hoạt động tốt cho các sản phẩm nhiều trang. Hình ảnh có định danh duy nhất ở chung cấp độ của các đối tượng (ví dụ một tài liệu, một bản ghi, một tập tin / thư mục, v.v), trong đó yêu cầu các thư mục được đặt tên theo mô tả. Các tên tập tin của hình ảnh là duy nhất trong mỗi thư mục con. Ví dụ, cuốn sách 0001 các tập tin hình 001.tif, 002.tif, và 003.tif. Sách 0002 chứa các tập tin hình ảnh 001.tif, 002.tif, và 003.tif. Sự nguy hiểm với cách tiếp cận này là nếu hình ảnh được tách ra từ thư mục cha của chúng thì sẽ không thể phân biệt hình ảnh đó thuộc thư mục nào.

## B.3 Các phiên bản

Vì những lý do khác nhau, một đối tượng quét duy nhất có thể có nhiều liên kết khác nhau với nó (ví dụ, cùng một hình ảnh nhưng chuẩn bị cho ý đồ khác nhau; phiên bản quét trên máy quét khác nhau, phiên bản có sửa đổi bổ sung; lớp, hoặc các kênh alpha được lưu trữ, quét từ phương tiện truyền thông ban đầu khác nhau, hoặc quét tại thời điểm khác nhau của các nhà cung cấp máy quét khác nhau). Lý tưởng nhất, mô tả và mục đích của các phiên bản khác nhau nên được đưa vào trong các Dữ liệu đặc tả. Nếu quy ước đặt tên phù hợp, các phiên bản phân biệt trong tên tập tin sẽ cho phép xác định nhanh chóng của một hình ảnh cụ thể. Một cách tiếp cận để đặt tên các phiên bản: thêm 001, 002 với tên tập tin cơ sở để chỉ phiên bản khác nhau là một lựa chọn; Tuy nhiên, nếu 001 và 002 đã biểu thị số trang, cần thực hiện cách hiển thị khác.

## B.4 Đặt tên tập tin phát sinh

Kế hoạch đặt tên tập tin cũng cần tính đến việc tạo ra các tập tin hình ảnh phát sinh được tạo ra từ các tập tin gốc trong quá trình thực hiện số hóa. Nói chung, tên tập tin phái sinh được thừa hưởng từ những tin gốc, thường là với một vòng loại vào để phân biệt vai trò của các phát sinh từ các tập tin khác (ví dụ như "p" cho phiên bản xuất bản, "l" cho hình thu nhỏ). Tập tin này thường bao hàm một sự thay đổi trong kích thước hình ảnh, độ phân giải hình ảnh, và / hoặc định dạng tập tin từ các tệp gốc. Tên tập tin phát sinh không cần phải được mô tả miễn là nó có thể liên kết trở lại các tập tin gốc. Đối với các tập tin phát sinh nhằm mục đích chủ yếu là để hiển thị trang web, cần phải xem xét cách đặt tên để có thể liên kết tới ảnh gốc chất lượng hơn, vì thế nó cần phải mô tả nhiều hơn để truy xuất được dễ dàng.

**Phụ lục C**  
(Tham khảo)

**Kiến nghị kiểm soát chất lượng**

Phụ lục này bao gồm một số lĩnh vực liên quan đến kiểm soát chất lượng, bao gồm hướng dẫn về kiểm tra, tiêu chuẩn và tần suất kiểm tra. Nó được thiết kế để hỗ trợ các tổ chức liên quan xây dựng của các quy trình kiểm soát chất lượng.

**C.1 Kiểm soát chất lượng hoạt động máy quét**

Máy quét phải được kiểm tra định kỳ để theo dõi hiệu suất hoạt động của nó và kiểm tra xem hình ảnh đầu ra có nằm trong dung sai thỏa thuận như các yêu cầu đặt ra theo các tiêu chuẩn hiện hành. Kết quả kiểm tra trước đó nên được sử dụng như một tiêu chuẩn cho hiệu năng hệ thống theo thời gian. Các biện pháp đơn giản đầu tiên đó là phải đảm bảo các thiết bị luôn được bảo dưỡng, giữ cho sạch sẽ.

**C.2 Xác nhận của đầu ra**

Các thiết bị nên thường xuyên ghi lại số lượng tài liệu lè và số lượng tài liệu hồ sơ (nhiều hơn một trang) được quét trong một phiên.

**C.3 Lấy mẫu**

Tần suất lấy mẫu để kiểm tra phải được xác định theo lịch trình nhất định. Tần suất này có thể được các nhà cung cấp hệ thống hỗ trợ và khuyến nghị

**C.4 Thiết lập bộ mẫu**

Thiết lập một mẫu các hồ sơ tài liệu cần số hóa nhằm mục đích đánh giá kết quả quét theo các tiêu chí chất lượng đã thỏa thuận. Những tài liệu này nên có đặc trưng chung cho các hồ sơ sẽ được quét và nên bao gồm các ví dụ về hồ sơ có chất lượng kém hơn so với phần lớn các tài liệu khác. Bảng kiểm tra mục tiêu và kiểm tra chi tiết được mô tả trong tiêu chuẩn ISO 12653-1 và ISO 12653-2, có thể được sử dụng để quét ảnh đen và trắng. Trường hợp số hóa liên quan đến hồ sơ có màu sắc, người thực hiện cần xem xét tiêu chuẩn màu sắc cho hình ảnh (chẳng hạn cho mục tiêu thương mại hóa và phải phù hợp với tiêu chuẩn ISO 12641.) Màu hiệu chuẩn, phù hợp với hồ sơ và phải được sử dụng nơi quét ảnh màu. Điều này có thể là một tiêu chuẩn thích hợp cho việc biểu diễn màu sắc hoàn hảo, tuy nhiên nó không phải là vấn đề quan trọng đối với độ phân giải và chất lượng tổng thể, là những tiêu chuẩn quan trọng và được yêu cầu trong ISO 29861.

**C.5 Các tiêu chí chất lượng cho hình ảnh**

Tiêu chí chất lượng cho hình ảnh nên bao gồm việc xem xét theo mức độ dễ đọc:

- chi tiết nhỏ nhất rõ ràng phải quét được (ví dụ như kích thước nhỏ nhất của chữ, sự rõ ràng của dấu chấm câu...);
- đầy đủ chi tiết (ví dụ các ký tự bị vỡ, mất dòng...);
- độ chính xác so với các bản ghi nguồn phi kỹ thuật số;
- máy quét tạo ra đốm (tức là đốm không có mặt trên bản gốc phi kỹ thuật số);
- đầy đủ của các phần của hình ảnh tổng thể (tức là thiếu thông tin ở bên các cạnh của hình ảnh);

- mật độ khu vực màu đen;
- độ trung thực màu sắc.

#### C.6 Dữ liệu đặc tả

- Thủ tục kiểm tra phải được thực hiện để đánh giá chất lượng Dữ liệu đặc tả được gán cho hình ảnh. Các vấn đề có thể được xem xét trong việc kiểm tra chất lượng của Dữ liệu đặc tả là:
  - tuân thủ các tiêu chuẩn của chính sách thể chế hoặc các yêu cầu của dự án số hóa;
  - thủ tục chứa hình ảnh với Dữ liệu đặc tả không đầy đủ;
  - liên quan và chính xác của dữ liệu đặc tả;
  - ngữ pháp - kiểm tra xem có đúng ngữ pháp, chính tả và dấu chấm câu, đặc biệt là dữ liệu bằng tay, có khóa; nhất quán trong việc tạo ra các dữ liệu đặc tả và trong việc chú thích của dữ liệu đặc tả;
  - đánh giá tính hữu ích của các dữ liệu đặc tả được thu thập;
  - đồng bộ hóa các dữ liệu đặc tả được lưu trữ tại các nơi, tức là phải được thực hiện để đảm bảo dữ liệu đặc tả được cập nhật một cách đồng bộ trên nhiều nơi (ví dụ như thông tin liên quan đến hình ảnh có thể được lưu trữ trong tiêu đề TIFF, hệ thống quản lý hồ sơ kỹ thuật số, và cơ sở dữ liệu khác);
  - tính đầy đủ của dữ liệu đặc tả - tất cả các điều kiện bắt buộc phải được thực hiện.

#### Đặc biệt quan trọng như sau:

##### a) Xác minh tính chính xác của các tệp tin định danh:

Tên tệp tin một cách nhất quán và định danh duy nhất cho cả hai nguồn tài nguyên kỹ thuật số và các bản ghi dữ liệu đặc tả (nếu nó tồn tại độc lập với hình ảnh). Tệp tin định danh có thể sẽ tồn tại cho các hồ sơ dữ liệu đặc tả riêng của mình ngoài việc định danh cho tài nguyên số hóa, có thể nhúng các thông tin như trang hoặc một phần số, ngày, dự án hoặc tổ chức nhận dạng. Thông tin nhúng trong tệp tin định danh cho tài liệu gốc phải được lưu trữ song song trong một bản ghi cơ sở dữ liệu hay tiêu đề. Định danh được xem như một liên kết tệp tin với thông tin lưu trữ trong cơ sở dữ liệu khác và phải mang lại chính xác dữ liệu đặc tả về tài liệu gốc. Xác minh định danh trên dữ liệu đặc tả tại các nơi khác nhau phải được thực hiện;

##### b) Thẩm định trình tự chính xác và đầy đủ của các tài liệu nhiều trang:

Trang phải theo đúng thứ tự không bị mất trang. Nếu các thành phần quan trọng của nguồn được ghi lại trong dữ liệu đặc tả, chẳng hạn như sự hiện diện của tệp đính kèm, tài liệu với các chương định nghĩa hoặc hồ sơ nhiều trang, thì chúng phải phù hợp với các tệp tin hình ảnh thực tế. Một quy ước về việc theo dõi và xem xét sự phù hợp với các tệp tin hình ảnh thực tế phải được mô tả.

## TCVN 12200:2018

### C.7 Tài liệu

Dữ liệu quản lý chất lượng (như nhật ký, báo cáo, quyết định) phải được lưu giữ trong hệ thống chính thức và phải trở thành một phần không thể thiếu của các dữ liệu đặc tả hình ảnh ở các tập tin hoặc cấp độ dự án. Những thông tin này có thể có giá trị lâu dài mà có thể có ảnh hưởng đến quyết định bảo tồn trong tương lai.

### C.8 Quy trình tái số hóa

Nếu có nhiều hơn 1% trong tổng số các hình ảnh và siêu dữ liệu được kiểm tra với một mẫu được lựa chọn ngẫu nhiên và không đảm bảo chất lượng với bất kỳ lý do được liệt kê ở trên, toàn bộ số lượng từ lần kiểm tra chất lượng cuối cùng sẽ được tái kiểm tra. Bất kỳ lỗi cụ thể được tìm thấy trong các mẫu ngẫu nhiên và bất kỳ lỗi nào khác được tìm thấy trong việc kiểm tra lại cần được sửa chữa. Nếu ít hơn 1% lỗi được tìm thấy, thì chỉ với những hình ảnh lỗi cụ thể và dữ liệu đặc tả sẽ được làm lại.

### C.9 Lỗi chung

Nói chung, lỗi chất lượng có thể được phân loại là "lỗi thực hiện", "lỗi quy trình" hoặc "lỗi điều hành". Lỗi thực hiện có thể tránh được bằng cách cung cấp thủ tục kiểm soát phù hợp để hướng dẫn việc số hóa. Lỗi quy trình thường nằm ngoài khỏi sự kiểm soát của các nhà điều hành và cần phải được giải quyết bởi một người giám sát quá trình này. Lỗi điều hành là những lỗi hàng ngày được thực hiện bởi các nhà điều hành khi họ làm việc.

#### C.9.1 Lỗi thực hiện

Có một số lỗi có thể tránh được với các đặc điểm kỹ thuật phù hợp. Chúng bao gồm:

- hồ sơ nguồn phi kỹ thuật số bị bản;
- không chính xác kích thước tập tin và định dạng, nơi các tập tin được thực hiện bị để kích thước sai hoặc lựa chọn sai lầm định dạng tập tin;
- nén, nơi các tập tin được thực hiện với một kiểu hoặc mức độ nén không phù hợp.

#### C.9.2 Lỗi quy trình

Có một loạt các lỗi quy trình có thể được gây ra bởi nhiều vấn đề nghiệp vụ. Những vấn đề này có thể bao gồm:

- tài liệu quy trình hoặc thông số kỹ thuật không đầy đủ hoặc không chính xác;
- bị lỗi phần cứng quét (hiệu chuẩn sai và các thiết bị đặc trưng);
- phần mềm bị lỗi (xử lý hình ảnh không chính xác hoặc bị lỗi liên kết hình ảnh trong cơ sở dữ liệu);
- hệ thống quản lý màu sắc thiết lập không chính xác;
- dữ liệu ban đầu chất lượng thấp;
- Dữ liệu đặc tả nguồn không chính xác.

### C.9.3 Lỗi điều hành

Các lỗi được gây ra bởi một số hình thức thực hiện trong công việc và có thể bao gồm:

- lỗi chụp cơ bản;
- cắt hình ảnh không đồng đều;
- hướng của hình ảnh bị sai, lộn ngược;
- tiếp xúc với hình ảnh quá sáng hoặc quá tối;
- tiêu cự không chuẩn;
- hiệu chuẩn hàng ngày, nơi mà các thiết bị chụp không được hiệu chuẩn;
- lỗi xử lý hình ảnh cơ bản;
- lỗi tối ưu hóa tập tin, trong đó thực hiện điều chỉnh không chính xác với màu sắc, độ tương phản và độ sáng của hình ảnh trong quá trình xử lý;
- đặt tên tập tin không chính xác, hoặc sử dụng tên không duy nhất;
- lỗi ghi dữ liệu đặc tả cơ bản;
- nhập dữ liệu vào hệ thống quản lý không chính xác, không đúng nơi dữ liệu cần được nhập;
- sử dụng từ vựng được kiểm soát không chính xác, sử dụng các từ không được xây dựng trong phạm vi ghi chú.

**Phụ lục D**  
(Tham khảo)  
**Các kỹ năng yêu cầu đối với nhân viên**

Phụ lục này mô tả các kỹ năng và nhiệm vụ mà nhân viên tham gia vào việc số hóa phải có. Điều này đảm bảo cho các tổ chức trong việc thực hiện tốt nhất quy trình số hóa.

**D.1 Khuyến nghị thiết lập kỹ năng**

Các kỹ năng cần được đào tạo cho cán bộ liên quan tới quy trình số hóa được mô tả trong bảng C.1.

**Bảng D.1 – Chức năng và mức độ kỹ năng được đề nghị**

<b>Bộ phận/linh vực</b>	<b>Yêu cầu</b>
<b>Quản lý</b>	<ul style="list-style-type: none"> <li>- đánh giá các trường hợp nghiệp vụ cần số hóa.</li> <li>- đàm phán mua, dịch vụ, bảo dưỡng thiết bị, vật tư.</li> <li>- kỹ năng quản lý nhà thầu số hóa hoặc giám sát của nhân viên thực hiện số hóa phụ thuộc vào mô hình số hóa tự làm hoặc thuê ngoài.</li> </ul>
<b>Phân tích nghiệp vụ</b>	<ul style="list-style-type: none"> <li>- xác định công việc cho quá trình số hóa.</li> <li>- xác định sự tích hợp của hồ sơ số hóa vào hệ thống thông tin doanh nghiệp / công việc hiện tại.</li> <li>- lựa chọn định dạng hình ảnh.</li> <li>- xác định yêu cầu nâng cao hình ảnh.</li> <li>- xác định kiến trúc thông tin để hỗ trợ quá trình kinh doanh.</li> </ul>
<b>Phân tích hệ thống</b>	<ul style="list-style-type: none"> <li>- lựa chọn phần cứng máy quét.</li> <li>- xác định yêu cầu lưu trữ.</li> <li>- tích hợp phần cứng, thiết bị hình ảnh và phần mềm.</li> <li>- tích hợp các yêu cầu số hóa vào cơ sở hạ tầng công nghệ thông tin hiện có của tổ chức.</li> <li>- tuân thủ các tiêu chuẩn công nghệ thông tin quốc gia và tổ chức.</li> <li>- thử nghiệm các cấu hình.</li> <li>- hỗ trợ sự liên tục của thiết bị số hóa (nếu cần thiết).</li> <li>- định nghĩa các chính sách và thủ tục để đảm bảo tính xác thực và tính toàn vẹn của hình ảnh kỹ thuật số.</li> </ul>

**Thư mục tài liệu tham khảo**

- [1] ANSI/NISO Z39.87-2006, Data dictionary - Technical metadata for digital still images.
- [2] ISO/TR 23081-1, Information and documentation - Records management processes - Metadata for records - Part 1: Principles.
- [3] ISO 15836:2009, Information technology - The Dublin core metadata element set (Công nghệ thông tin - Tập phần tử Dữ liệu đặc tả Dublin core).
- [4] ISO 19115-1:2014 *Geographic information - Metadata - Part 1: Fundamentals*.
- [5] TCVN 6909:2001 (ISO/IEC 10646-1:2000) Công nghệ thông tin - Bộ mã ký tự Tiếng Việt 16-bit.
-