S-101PT12 - 13-15 February 2024 – VTC

Attribute formats and constraints in S-101

Agenda Item S-101 PT12-06.12

Example:

27.202 waterway distance

IHO Definition: **WATERWAY DISTANCE**. The length of the space between two points along a waterway. (Adapted from Oxford English Dictionary).

Attribute Type: Real

Unit: Defined by the sub-attribute **distance unit of measurement** (see clause 27.91).

Resolution: 0·1

Format: xx.x

Example: **2.5** for a waterway distance value of 2.5 nautical miles (where **distance unit of measurement** is populated as *5* (nautical mile).

- ❖ The format description may lead the encoder think that the Integer part of the attribute value must have no more than 2 digits;

- ❖ Not sufficient for a distance that would be encoded in metres with a value ≥ 100;

- ❖ This sometimes leads to difficulties in encoding validation checks;

- ❖ Similarly, resolution 0.1 would not be sufficient for a distance that would be 5.25 nautical mile.

Attribute Types

> **S-100 defines 13 attribute types, 7 of them are retained in S-101**

S-100 - 10a-5.1.4		S-101 DCEG - 2.4.2
S-100 Attribute Type	**Value Type**	**Attribute type**
S100_GF_TextAttributeType	CharacterString	Text (TE)
S100_GF_IntegerAttribute	Integer	Integer (IN)
S100_GF_RealAttribute	Real	Real (RE)
S100_BooleanAttributeType	Boolean	Boolean (BO)
S100_EnumerationAttributeType	PositiveInteger	Enumeration (EN)
S100_DateAttributeType	Date	Not Used
S100_TimeAttributeType	Time	Time (TI)
S100_DateTimeAttributeType	DateTime	Not Used
S100_URIAttributeType	URI	Not Used
S100_URLAttributeType	URL	Not Used
S100_URN_AttributeType	URN	Not Used
S100_TruncatedDateAttributeType	S100_TruncatedDate	TruncatedDate (TD)
S100_CodelListAttributeType	CharacterString	Not Used

Note: in Edition 1.2.0 of the DCEG, Simple attribute types "Free Text" (TE) has been amended to "Text"

in 2.4.2. This is to be aligned with in section 27 to 30.

Attribute Format

> **S-100 provides guidance for the format of each attribute type. Ex:**

S100_GF_RealAttribute [Real]	Decimal floating-point numbers in the value domain of double precision numbers (IEEE 64-Bits).	The decimal separator is always a point ('.') and there must be no thousands separator used. Positive numbers should not use the '+' sign. Negative numbers must have a leading '-' sign. The exponential form is permitted. Non-significant zeros are prohibited. If there are only zeros to the right of the decimal point the decimal point should be omitted. The values INF, -INF, and NaN are prohibited.	123.456 -42 1E-5 -2.45E7

DCEG 1.3.1: "NOTE: *A feature attribute type has a <u>name</u>, a <u>data type</u> and a <u>domain</u> associated to it***".**

"Format" as listed in the DCEG does not conform to any formal description. The information supplied by the <u>attribute type</u> and the <u>Precision</u> (Resolution) are self-sufficient.

27.191 waterway distance

Waterway distance: <u>IHO Definition:</u> The length of the space between two points along a waterway. (Adapted from Oxford English Dictionary).

<mark>Attribute Type: Real</mark>

<u>Unit:</u> Defined by the sub-attribute **distance unit of measurement** (see clause 27.89).

<u>Resolution:</u> 0·1

<mark>Format: xx.x</mark>

It is proposed to remove references to "Format" in sections 27 to 30 of the DCEG.

Attribute Constraints (1/3)

> **S-100 defines possible constraints on attributes:**

2a-4.2.10 S100_CD_AttributeConstraints

Role Name	Name	Description	Mult	Type	Remarks
Class	S100_CD_AttributeConstraints	Constraints of a simple attribute	-	-	
Attribute	stringLength	Shall be represented as a positive integer (that is, greater than zero) that specifies the maximum number of characters that may be assigned to the text attribute type. If not specified, then the text length shall be unconstrained	0..1	PositiveInteger	
Attribute	textPattern	A character string that specifies a scheme of one or more constraints on the structure of the text values that may be assigned to the attribute. This shall be achieved by using a regular expression. W3C XML Schema Part 2: Datasets Second Edition, Appendix F (Regular Expressions) shall be used to define text patterns in this standard	0..1	CharacterString	
Attribute	range	Specifies the range of allowed numeric values	0..1	S100_NumericRange	
Attribute	precision	Specifies the precision of a real number	0..1	NonNegativeInteger	

Note: the range can be defined by Lower (Minimum) and Upper (Maximum) values + Closure

IHO

International Hydrographic Organization

> **Constraints in the Feature Catalogue**

Although S-100, 2a-3.5 states that attribute constraints can be specified in the FC, very few instances are present in the 1.2.0 S-101 FC. Ex:

```
<S100FC:S100_FC_SimpleAttribute>
    <S100FC:name>Text Offset Bearing</S100FC:name>
    <S100FC:definition>The angular distance measured from true no
    <S100FC:code>textOffsetBearing</S100FC:code>
    <S100FC:valueType>integer</S100FC:valueType>
    <S100FC:uom>
        <S100Base:name>Degree</S100Base:name>
        <S100Base:symbol>°</S100Base:symbol>
    </S100FC:uom>
    <S100FC:constraints>
        <S100CD:textPattern>xxx</S100CD:textPattern>
        <S100CD:range>
            <S100Base:lowerBound>0</S100Base:lowerBound>
            <S100Base:upperBound>360</S100Base:upperBound>
            <S100Base:closure>closedInterval</S100Base:closure>
        </S100CD:range>
    </S100FC:constraints>
</S100FC:S100_FC_SimpleAttribute>
```

The (machine readabale) FC should be the basis for attribute constraints (the DCEG being the « human readable version).

It is proposed to systematically populate attribute constraints in the S-101 FC (requires a systematic review).

This includes attribute range and the use of Regular Expressions (regex) for Text attributes (when needed).

S-101 PT12-06.12 – Attribute Format and Constraints in S-101

Attribute Constraints (3/3)

➢ **Alignment of S-100 and S-101 terminology**

The table below summarize the proposed changes (in red) for attribute constraints in S-101 2.0.0.

S-100 - 5.1.0		S-101 - 1.2.0		S-101 - 2.0.0	
S-100 - 2a-4.2.10	S-100 - 13-8.1.1.1	FC	DCEG	FC	DCEG
stringLength	stringLength	stringLength		stringLength	Remark
textPattern	textPattern	textPattern		textPattern	Remark
range	rangeLower	lowerBound	Minimum value	lowerBound	Minimum value
	rangeUpper	upperBound	Maximum value	upperBound	Maximum value
	rangeClosure	closure		closure	
precision	precision	precision	Resolution	precision	Precision
		/	Format	/	/

It is proposed to:

- replace "Resolution" by "Precision" in the DCEG;

- add guidance in 2.4.2 explaining that "Precision" indicates the number of decimal digits.

S-101 PT12-06.12 – Attribute Format and Constraints in S-101

Inconcistencies (1/2)

> **Various inconcistencies have been identified in the IHO documentation**

<u>Between FC and DCEG:</u>

- ✓ For attribute source, stringLength constraint is "150" in the FC, whereas there is no guidance in the DCEG; textPattern constraint is "c…" where there should be no constraint;

- ✓ For attribute sector bearing, textPattern constraint is "xxx.xx" in the FC. This should be replaced by precision = 2;

- ✓ For attribute depth range minimum value, textPattern constraint is "sxxxxx.xx; s = sign, negative values only" in the FC. This is useless as already specified by the attribute type (Real).

```xml
<S100FC:S100_FC_SimpleAttribute>
    <S100FC:name>Source</S100FC:name>
    <S100FC:definition>The publication, document, or r
    <S100FC:code>source</S100FC:code>
    <S100FC:remarks>May be populated with the correspo
    <S100FC:definitionReference>
        <S100FC:sourceIdentifier>220</S100FC:sourceIde
        <S100FC:definitionSource ref="IHOREG" />
    </S100FC:definitionReference>
    <S100FC:valueType>text</S100FC:valueType>
    <S100FC:constraints>
        <S100CD:stringLength>150</S100CD:stringLength>
        <S100CD:textPattern>c…</S100CD:textPattern>
    </S100FC:constraints>
</S100FC:S100_FC_SimpleAttribute>
```

S-101 PT12-06.12 – Attribute Format and Constraints in S-101

Inconcistencies (2/2)

Name	Orientation Value
CamelCase	orientationValue
Item Identifier	860 ?
Definition	The angular distance measured from true north
Data type	real

	Minimum Range	0.00
	Maximum Range	360.00
	Range Closure	0.00 = 360.00
Additional Data	Precision	0.01
	Quantity Specification	planeAngle
	Format	
	Definition	
	Symbol	

Between the S-101 FC and the IHO Registry

- ✓ Attribute "orientation value" has a range closure: "0.00 = 360.00" which does not conform with S-100 closure constraint (should be "geLtInterval" or "[0,360)");

- ✓ Attribute precision is listed as (for example) "0.01" where it would be "2" (number of decimal digits) in the FC. Note: S-100, defines Precision as a non-negative Integer;

- ✓ GI Registry provide indication for "Format" (which seems useless), but no "Unit" (which is essential) for simple attribute.

It is proposed to review (out of scope of the S-101PT) the presentation of IHO GI Registry to have it consistent with how attributes and constraints are defined in S-100.

S-101 PT12-06.12 – Attribute Format and Constraints in S-101

Recommendations

A. Systematically encode the attribute constraints (including regex) in the S-101 Feature Catalogue;

B. Review attribute guidance in the DCEG, (sections 27 to 30):

   B1: remove reference to "Format";

   B2: replace "Resolution" by "Precision" and add guidance in 2.4.2;

   B3: amend "Free Text" to "Text" in sections 27 to 30;

   B4: ensure the guidance on the constraints will be consistent with those in the FC.

C. Review (out of scope of the S-101PT?) the presentation of IHO GI Registry to have it consistent with how attributes and constraints are defined in S-100.

QUESTIONS?

THANK YOU