

Title: Ensure consistent usage of ISO 639-2 3-letter language codes

S-100 Maintenance - Change Proposal Form

Organisation	NIWC	Date	7/7/2021
Contact	David Grant	Email	David.Grant1@navy.mil

Change Proposal Type *(Select only one option)*

1. Clarification	2. Correction	3. Extension
	X	

Location *(Identify all change proposal locations)*

S-100 Version No.	Part No.	Section No.	Proposal Summary
5.0 20210330 draft	10b GitHub Schema	9.6.1 S100XSLTPR	Update <i>datasetLanguage</i> Replace xs:language with string restriction

Change Proposal

Ensure consistent usage of three letter language codes. Details attached.

Change Proposal Justification

S-100 language codes must use an ISO 639-2 3-letter code.

The S-100 GML encoding currently specifies use of an ISO 639-1 two letter code ("EN").

The portrayal and alert schema specify that xs:language is used to store language codes. Although xs:language permits three character codes, it also allows two character codes. RFC 3066 says for languages with both two and three letter codes the two-letter code must be used. xs:language also allows various sub-tags to be added using hyphens.

These are all valid per xs:language, but only "mas" conforms to S-100:

Code	Language	Explanation
en	English	ISO-639 two-letter language code
mas	Masai	ISO-639 three-letter language code
fr-CA	French as used in Canada	ISO-639 two-letter code with ISO-3166 two-letter country code
en-scouse	English Liverpudlian dialect known as 'Scouse'	ISO-639 two-letter language code with addition, IANA-registered
i-klingon	Klingon	IANA-registered language code
x-pig-latin	Pig Latin	Unregistered/Experimental

What parts of the S-100 Infrastructure will this proposal affect?

- ☐ S-100 Feature Concept Dictionary Interface or Database
- ☐ S-100 Portrayal Register
- ☐ S-100 Feature Catalogue Builder
- ☐ S-100 Portrayal Catalogue Builder
- ☐ S-100 UML Models
- ☒ S-100 GitHub Schemas

Please send completed forms and supporting documentation to the secretary S-100WG.

10b-9.6.1 Dataset identification

Table 10b-3 – Dataset identification header elements

Field	XML Tag	Value	Mult.	Type	Description
...
Dataset language	datasetLanguage	“engEN”	1	ISO 639-2 3-letter code ⁴	The (primary) language used in this dataset
...

S100CSL.xsd

```

...
<!-- UnlimitedInteger -->
<xs:simpleType name="UnlimitedInteger">
  ...
</xs:simpleType>
...
<!-- ISO 639-2 3-letter language code -->
<xs:simpleType name="languageCode">
  <xs:restriction base="xs:string">
    <xs:pattern value="[a-z][a-z][a-z]"/>
  </xs:restriction>
</xs:simpleType>

```

S100PortrayalCatalog.xsd

```

...
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:s100Symbol="http://www.iho.int/S100SymbolDefinition" version="4.1.0-20191115">
  xmlns:s100CSL="http://www.iho.int/S100ConceptualSchema"
  <xs:import namespace="http://www.iho.int/S100SymbolDefinition"
    schemaLocation="S100SymbolDefinition.xsd"/>
  ...
  <!-- Class for descriptive information about a catalogue item -->
  <xs:complexType name="Description">
    <xs:sequence>
      <xs:element name="name" type="xs:string" minOccurs="0" maxOccurs="1"/>
      <xs:element name="description" type="xs:string"/>
      <xs:element name="language" type="s100CSL:languageCode"/>
    </xs:sequence>
  </xs:complexType>

```

S100AlertCatalog.xsd

```

...
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:s100Symbol="http://www.iho.int/S100SymbolDefinition">
  xmlns:s100CSL="http://www.iho.int/S100ConceptualSchema"
  <xs:include schemaLocation="S100PortrayalCatalog.xsd"/>

```

```

    <xs:import namespace="http://www.iho.int/S100SymbolDefinition"
schemaLocation="S100SymbolDefinition.xsd"/>
<!-- MESSAGES -->

<xs:complexType name="Text">
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="language" type="xs:languageS100CSL:languageCode"
default="eng"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
...

```