Paper for Consideration by TSMAD27

S-100 Metadata Schema Progress Part 3

Submitted by:	TSMAD Vice Chair
Executive Summary:	This paper presents a further update to the S-100 metadata schema progress
Related Documents:	TSMAD25 paper 4.3.2 TSMAD26 paper 11.4A
Related Projects:	S-100, S-101

Introduction / Background

At TSMAD25, the UK presented a paper that identified the need for metadata schemas for S-100. The work was further updated at TSMAD26. The UK and IIC were tasked with further developing the appropriate metadata schema's and this paper represents these results and recommendations.

Analysis/Discussion

In order to develop the appropriate S-100 metadata schemas, IIC reviewed the originating ISO documents and found that Annex C in ISO 19115 and sections A3/A4 in ISO 19139 outline the rules for creating various metadata extensions and community profiles.

In addition, IIC reviewed Schematron and Dutch and ANZLIC (Australia New Zealand Spatial Information Council) metadata Schematron validation rules in order to develop the appropriate S-100 metadata schema.

For S-100, TSMAD is imposing a more stringent obligation on an existing fileIdentifier metadata element. This is shown via theS-100 UML diagrams that have been derived from ISO 19115 and enforced via a tool other than an XML schema validator. IIC found that Schematron seems to be a good way to enforce such obligations. It was used by ANZLIC to enforce identical restriction in their metadata profile and it is also used extensively (100+ tests) to validate Dutch metadata instance documents. It is worth noting that while the S100_Metadata class specializes the MD_Metadata class, the specialization only involves restricting fileIdentifier from optional to mandatory. As such, in order to ensure interoperability with ISO standards the MD_Metadata root element should be used for S-100 XML implementations. This will ensure that all tools capable of reading ISO 19115 based metadata XML instances will also be able to read S-100 based dataset metadata XML instances without any modifications required. This is also what ANZLIC has done.

In line with the above, IIC created two things: a sample XML file illustrating what S-100 minimum dataset metadata (as outlined in section 4a-5.4) should look like and the Schematron rule file for validation of constrains for fileIdentifier element. Two tests were implemented, one for making sure fileIdentifier exists in the metadata file and another one to ensure it is not empty (after removing all white space).

NOTE: The XML file will have to be generalized for S-100 as it currently uses NOAA data for validation.

IIC has implemented the above in the attached Schematron rule file (plus the regular ISO 19139 schema) and performed some simple testing using the sample XML minimum metadata file.

Recommendations

As part of their review IIC also recommended that S-100 part 4 should also re-write 4A-C1.1 (paragraphs 2-4) accordingly, perhaps as follows (please feel free to edit as required):

> *It is necessary to implement the Profile in order to prove compliance. ISO/TS 19139:2007 is an XML schema implementation of ISO 19115:2003 and can be used to prove partial compliance to both ISO 19115:2003 and S-100 Metadata Profile. IHO has developed an additional Schematron rules to enforce the additional restriction for fileIdentifier element. Consequently, proof of compliance to the S-100 Metadata Profile will be via validation of the XML document instances against the ISO/TS 19139:2007 XML Schema Definition (XSDs) and S-100 Schematron Metadata Rules.*

This re-write will be part of a formal S-100 proposal to fix other inconsistencies in the S-100 metadata profile.

Action Required of TSMAD

The TSMAD is invited to:

a. endorse the work of IIC regarding the S-100 metadata schemas
b. agree to providing these metadata schema's as part of S-100
c. note that there will be a proposal regarding other aspects of S-100 metadata.