S100WG01 8.11A

Paper for consideration by S-100 WG

Comments on Redline for Part 4A Metadata

Submitted by:	Jeppesen
Executive Summary:	Input on proposed amendments to S-100 Part 4A.
Related Documents:	TSMAD29/DIPWG7-11.4A.
Related Projects:	(1) S-100 2.1.0

1 Introduction/Background

This document contains input to the proposals for updating S-100 Part 4A circulated in January 2016 prior to S-100WG-1. The main part of this paper summarizes the more significant comments. More detailed comments are in the attached appendix and the working group is requested to review and comment on that list as well.

2 Analysis/Discussion

2.1 Updating S-100 Metadata for ISO 19115-3

In 2014, the ISO replaced ISO 19115 with a revised standard, ISO 19115-1. The XML schema implementation will be released in a new technical standard ISO 19115-3. This will presumably replace ISO 19139. As of the writing of this paper, ISO 19115-3 was in "Draft" status and had not been released by ISO. (IHB liaison queried the ISO and it is awaiting release.) The XML schemas corresponding to ISO 19115-1 are therefore formally still in draft status and had not yet been placed on ISO's official site for publicly accessible standards, though draft versions are accessible from other locations. The UML models for ISO 19115-1 are in that document and have therefore been formally released. (TSMAD29/DIPWG7-11.4A summarizes the changes to ISO 19115-1.)

We believe it is better to harmonize the update to S-100 Metadata with an updated XML schema implementation, instead of updating S-100 and then perhaps having to revise it after ISO 19115-3 is released.

2.2 Reuse of ISO 19115 types

The ISO 19115 types CI_Citation and EX_Extent are similar to "S100_" classes defined in Part 4A and the model will be simplified by re-using the ISO type CI_Citation in place of S100_ProductSpecification and S100_SupportFileSpecification. Consideration should also be given to using the ISO type EX_Extent in bounding box and data coverage.

CI_Citation attributes title, edition, and editionDate are functionally substitutes for the name, version, and date attributes of the specification classes. The constraints on attribute multiplicity can be checked with Schematron rules.

EX_Extent has a CharacterString attribute "description" and a component EX_GeographicExtent (sub-types EX_GeographicBoundingBox and EX_BoundingPolygon) which combine the functions of the ID, boundingBox, and boundingPolygon attributes of S100_DataCoverage.

2.3 Other comments

The process for using checksums is less complex than digitally signing something and checksums might still have a role to play. They should be restored and made optional.

The question of digital signatures in catalogue metadata needs explanations. Questions that arise include: What is being signed here, why, and who will sign it? If it is intended for signing catalogues – will it be the IHB or each producing agency? Will we end up with all producers having to sign all portrayal and feature catalogues? Will a system end up with multiple copies of catalogues differing only in who signed them? If some need to be signed, does signing have to be mandatory for all?

3 Recommendations

Recommendations are detailed in the appendix.

4 Impacts

Few at present, since the subject is a proposal in development, but generally speaking simplification of the model and more re-use of ISO types.

5 Actions Requested

The S-100WG is invited to:-

- discuss and comment on the recommended changes to the draft of Part 4A.

Annex A. Detailed Comments

Clause / sub-clause / annex	Paragraph / Figure / Table / Note	Type	Comment – justification for change	Proposed change	Observations
all		te	Needs to be updated in conformance with ISO 19115-1:2014. This action would ideally be synchronized with the XML schema implementation of metadata, ISO/TS 19115-3 but the latter has not yet been released by ISO. As of January 12 2016 it is in DTS stage and a mature version of the schemas is available though not the text.	Depends on the status and ISO plans for ISO 19115-3. Alternatives: 1) Proceed with updating the text (only) of S-100 Part 4A. Update the XML schema implementation ASAP when ISO 19115-3 is released. 2) Hold back on updating the text of S-100 Parts 4A-C until ISO 19115-3 is released.	
Overview	Figure 4A-D.4	te	ISO 19115 already defines a citation type CI_Citation type which can be used instead of S100_SupportFileSpecification and S100_ProductSpecification. CI_Citation has attributes title, edition, and editionDate as substitutes for name/version/date	1) Delete S100_SupportFileProductSpecification and S100_ProductSpecification classes 2) Change the 3 attributes using them as follows: S100_DatasetDiscoveryMetadata.productSpecification : CI_Citation [1] S100_ExchangeCatalogue.productSpecification : CI_Citation [0..1] S100_SupportFileDiscoveryMetadata.supportFileSpecification [0..1] (mandatory if and only if dataType value is not Text). 3) Add attribute locale: PT_Locale [0..1] or characterSet: MD_CharacterSetCode [0..1] to S100_SupportFileDiscoveryMetadata for the character encoding.	

Overview	Figure 4A-D.4	te	S100_DatasetDiscoveryMetadata.dataCoverage and S100_DataCoverage.boundingBox are both mandatory in this diagram, implying that a geographic bounding box is required in metadata. But there may be S-100 products where a bounding box is not defined, e.g., “Norwegian ports”. In the original, bounding box/polygon were optional. ISO 19115 defines dataType EX_Extent which allows use of EX_BoundingPolygon, EX_GeographicBoundingBox and a text attribute called “description”.	Add attribute extent: EX_Extent [0..*] to S100_DatasetDiscoveryMetadata for the dataset bounding box. Make dataCoverage optional (mult. 0..*) in S100_DatasetDiscoveryMetadata. Change S100_DataCoverage replacing attributes boundingBox & boundingPolygon with attribute extent: EX_Extent [0..*] with a constraint requiring (either?) EX_GeographicBoundingBox or EX_BoundingPolygon. (Also, extent.description can replace ID.)	
S100_SupportFileDiscoveryMetadata		te	checksum and digital signatures are different, a file might have a checksum but not be signed	restore the checksum attribute but make it optional (multiplicity 0..1)	
S100_SupportFileFormat	literal ASCII	te	S-100 text is in UTF8, not ASCII	Revert literal from ASCII to Text	
S100_CatalogueMetadata	attributes	te	Why do fileName, fileLocation, versionNumber, issueDate, and productSpecification have multiplicities of 1..* instead of 1, 0..1? It will need some complex constraints to match the numbers of these attributes. If it is to allow for multiple catalogues, an upper bound of * on the role multiplicity in figure 4a-D.3 should be able to handle it. Alternatively, a complex attribute to group the info for each catalogue should be introduced.	multiplicity of all attributes should be [1] or [0..1] multiplicity at appropriate association end in figure 4a-D.3 should be 1..* or 0..*	
S100_CatalogueMetadata	productSpecification	te	CI_Citation can be used instead	change type to CI_Citation	
S100_CatalogueMetadata		te	catalogues might be in different languages and character sets	Add as attributes either: locale: PT_Locale OR, both of: language: ISO 639-2 & characterSet: MD_CharacterSetCode.	

S100_CatalogueMetadata	digitalSignatureReference & digitalSignatureValue	te	What is being signed here, why, and who will sign it? If for signing catalogues – will it be the IHB or each producing agency? Will we end up with all producers having to sign all portrayal and feature catalogues? Will a system end up with multiple copies of catalogues differing only in who signed them? If some need to be signed, does signing have to be mandatory for all? At present it is not clear why digital signatures are needed here, the same is goes for the implications of requiring signatures here. Checksums might suffice to guard against transmission or other corruption.	The remarks column should describe what is being signed. Analyse the implications of signing and explain them in a new clause. Make digitalSignatureReference and digitalSignaturevalue optional.	