TSMAD 25-4.3.2

Paper for consideration by TSMAD

Revisions and Extensions to S-100 Edition 1.0.0

Submitted by:	Jeppesen/UKHO
Executive Summary:	Outlines some revisions and extensions to S-100 Edition 1
Related Documents:	(1) S-100 Ed. 1.0.0
Related Projects:	(1) S-100

Introduction / Background

The S-100 standard describes a framework for modeling and exchange of maritime geospatial data intended to support a wide variety of data sources, products, and customers. After the issue of S-100 Edition 1.0.0 in 2010, several groups have been working on developing Product Specifications for different domains and applications. This work has identified missing elements and ambiguities in S-100 Edition 1.0.0 as well as some extensions that are needed for domains of maritime data other than hydrographic data. This paper outlines the missing elements, ambiguities, and necessary extensions we have identified so far.

This paper is not intended to be a comprehensive review of the IHO standard S-100 Edition 1.0.0, but instead intended to identify discrete areas of the framework where updates to S-100 are needed.

Terms and Abbreviations

FAL (Convention on) Facilitation of International Maritime Traffic

References

CEN 15449: Geographic information — Standards, specifications, technical reports and guidelines, required to implement Spatial Data Infrastructures. BSI PD CEN/TR 15449:2011.

INSPIRE D.2.5: INSPIRE Generic Conceptual Model. 2012-05-15. INSPIRE Project document D2.5_v3.4rc2

ISO 19109: Geographic Information – Rules for Application Schema.

ISO 19757-3: Information technology Document Schema Definition Languages (DSDL) Part 3: Rule-based validation Schematron - First Edition.

S-100: Universal Hydrographic Data Model, Edition 1.0.0, January 2010.

BLAST1: Harmonization of Nautical Information. Technical report for BLAST project, D_WP4_3, 2011

Parts 2 (Management of Registers) & 2a (Feature Concept Dictionary Registers)

Reconciliation of S-100 and S-99

Parts 2 and 2a need to be updated and re-structured to take into account the existence and contents of IHO Publication S-99.

Descriptions of operational procedures are given in both S-99 and Parts 2 and 2a of S-100. S-100 also specifies the schema for an IHO Geospatial Information Register (§ 2-7) and Feature Concept Dictionary (Part 2a).

There will be changes to S-99, following recent discussions between IHO and IALA working groups on the question of main/supplementary registers.

Operational procedures given multiple domains

The working of operational procedures in the presence different domains and control bodies may need clarification. For example, it may be necessary to transfer a concept between domains (perhaps on creation of a new domain). Operational processes currently described in S-99 should be able to handle the possible situations, but their working in scenarios likely in the presence of multiple domains should be clarified. Some possible scenarios are:

- Transferring a single concept from one domain to another.
- Merging two or more domains.

We are not suggesting that procedures be defined for every possible situation, but an evaluation of what are the likely scenarios and how efficiently current procedures would work should be undertaken. An alternate solution may be to address these special changes as necessary in consultation with the registry manager.

Recommendations:

1) Ensure that the normative parts are in just one document instead of both S-99 and S-100. Specifically, re-structure S-100 Parts 2/2a as:
    a. A non-normative overview of registry organization and content, which refers to S-99 as the normative reference for organization and procedures.
    b. Normative sub-parts 2a, 2b, etc., specifying the register schema (currently § 2-7); Feature concept dictionary registers (currently Part 2a); and other registers (i.e., portrayal, metadata, product specifications, and data producer codes registers).
2) Evaluate the operational procedures already defined for applicability and efficiency in the presence of multiple active domains.

Part 3 (General Feature Model and Rules for Application Schemas)

Codelists and Enumerations

The underlying ISO standards (including ISO 19136) make extensive use of Codelists, but S-100 classifies ListValue as an enumeration (1-4.8.1 (c)). This makes no difference for ISO8211 encodings but in XML has to be turned into simpleType definitions.

The INSPIRE Generic Conceptual Model (2012-06-15) recommends that:

- Enumerations should be used if the set of allowed values is fixed, and code lists should be used if the set of allowed values may be extended without a major revision of the data specification.
- For code lists, if there is an existing list managed by an international organization, this code list should be referenced.
- If the existing code list is not continuously available elsewhere, it may be made available through the common INSPIRE code list registry.

S-100 should also take into account the existence of multiple user communities (charting, notices, AIS, etc.) which may use the same sets of allowed values.

Adopting these literally might require that each application schema define an enumeration for such well-known lists as time zones, ISO country codes and ISO language codes.

Recommendations

1) Add an attribute type Codelist and adopt the INSPIRE recommendations as to codelist vs. enumeration.
2) Define a code list register and define operational procedures for maintaining the code list register and individual code lists (the latter only if a permanent location maintained by a responsible organization cannot be located).

3) Use code lists for ISO or non-ISO widely-used[1] codes: country, language, horizontal or vertical datum or CRS, time zones, and ENC scale bands(?). (For ISO 8211 encodings, numeric codes might have to be defined if not already available.)

4) Use enumerations for other allowed values.

Allowing domain-specific models for non-geospatial information

Given that the S-100 will also be used as a framework for domains other than ENCs, it may be necessary to revise it to make it more expressive.

Some forms of data may be easier to model using approaches other than the object-attribute approach. For example, trying to use a purely object-attribute approach to modelling rules (e.g., country-specific navigation regulations) is complicated for anything more than the simplest rules. S-100 exchange sets should provide for domain-specific modelling and encoding specifications of information that is primarily non-geospatial and hence not a good fit for the underlying ISO 191xx family of standards, but which still needs to be part of a maritime information transfer set or transactional data interchange. E.g., a "message",

Example 1: National shipping regulations

Sailing directions tend to contain regulations expressed in relatively long paragraphs, often reproduced from the official legal publication. NHS, BSH, and KMS attempted to map examples of the shipping regulations for Norway, Germany, and Denmark to an object-attribute model as part of the BLAST project. Their experiences are reported in a BLAST technical report [BLAST1]. All experienced difficulties converting their national shipping regulations into the object-attribute framework. This experience shows that converting IF-THEN statements, often with multiple conditions, sub-clauses and exceptions, into an object-attribute model is a complex activity which is susceptible to error, and it is sometimes not possible to capture important details of the original information.

Example 2: IVEF information delivery

Inter-VTS Exchange Format being defined by the IALA e-navigation committee working group and the IALA VTS committee defines information exchange by means of a web service, with authentication, query, and response. This is different from the "transfer set" encapsulation; instead it needs the notions of "data stream" and "data packet" for data delivery and "transaction" as the mechanism of data transfer.

Example 3: AIS messages

These messages (with source, message header, body, etc., and of type broadcast or addressed) rather than exchange sets.

Recommendations

1) S-100 should support "exchange" models in addition to the "transfer set". Some such models are service-centric views (Ref. CEN 15449, ISO 19119, WFS, WMS, WSDL, SOAP).

2) Product specifications should use "support files" to define domain-specific models where such are needed for domain-specific information. S-100 should allow product specifications to extend the notion of an exchange set, dataset file or support files to include individual files or collections of files within constraints to be decided. (This recommendation should result in minor changes at most to S-100 itself, mainly in setting the abovementioned constraints.) Consider whether this might lead to extensions of S-100 which might affect compatibility with other S-100 machines.

3) A short thought experiment in adding a domain-specific model to an S-100 product should help identify whether and how S-100 needs to be adapted. An example showing how to define domain-specific models can be a non-normative Annex or Part of the base S-100 standard.

GFM Revisions

Updated GFM

The GFM needs to be updated to include changes made since Edition 1.0.0.

[1] Possible guidelines/rules-of-thumb for "widely-used" are: (a) whether they are likely to be used in 3 or more domains and (b) likely to use the same list of allowed values in those domains of use.

Recommendations

1) Changes to the GFM and Feature Catalogue schemas to improve the expressiveness of association modelling were agreed at TSMAD22/DIPWG3 and remain to be incorporated into S-100.

2) Changes to the GFM and Feature Catalogue to accommodate the use of numericCode (for use in ISO 8211 encodings) and alias remain to be discussed.

URI schemes and universal identifiers

The issues of URI schemes and universal identifiers should be discussed in the context of an update or new edition of S-100. What items, if any, should be denoted by URIs? Feature classes, attributes, allowed values? A separate paper on identifiers and feature register will be submitted by Jeppesen.

Roles

As we understand the (updated) GFM, a "role" is mandatory at both ends of feature associations and one end of information associations (the role at the other being fixed). Often, specifying the role adds little to the information content[2] of the data but enlarges the feature catalogue and increases the data volume. Unless a compelling case exists for requiring the dataset to contain the roles can be made from either the modeling or implementation perspective, the specification of roles should be optional. For cases where the data need to distinguish source and target, identification as "source" or "target" (or forward/inverse) should be sufficient.

Recommendations:

1) Make roles optional.

2) Define default roles "source"/"target" (add "both" for the as-yet-unknown case where the domain model must treat both the same).

3) Use signed association codes to denote the direction in the encoding.

Part 4a (Metadata)

S-100 Metadata schemas yet to be made available

The GML/XML document instances mentioned in Appendix 4A-C §C-1.1:

> *The XML documents provided by IHO allow full implementation of ISO 19115:2005. IHO has also developed XML documents to allow for the implementation of the S-100 Metadata Profile. The ISO/TS 19139:2007 XSDs have been used for this implementation.*
>
> *The XML documents consist of:*
>
> > *1) ISO/TS 19139:2007 XSDs,*
> >
> > *2) GML / XML document instances for each of the S-100 Geographic Extent Name category lists and the S-100 Search words so they can be registered according to the ISO 19135:2006 standard and be referenced from XML metadata document instances.*

[The metadata XML schema defined for marine protected areas is offered to TSMAD as a starting point.]

Metadata Definitions

Generally speaking, the metadata definitions in S-100 should reflect its status as a framework standard by being defined so that they can be used for domains other than ENCs. TSMAD is invited to consider whether elements appropriate to ENCs are also appropriate or even available for other domains.

An example of changes are needed for other domains is making bounding box optional, which is appropriate for notices of arrival/departure. These would otherwise be encoded as geographic areas given by the territorial waters of a state, or a point coverage dataset of a state's international ports, or even perhaps world-wide coverage. Ideally, data coverage would be described without numeric

[2] The *name* of the association is often sufficient information.

coordinates in a notice of arrival/departure exchange set (i.e., in terms of “country X” or “Ports of Country X”). Such an exchange set is discoverable by textual search for the name of a country though not by spatial query/browse.

There may be a need to separate product or service and message metadata as a very small non-19115 subset will be appropriate for some message type data.

Recommendations

1) Provide the metadata schemas referenced in S-100.
2) Review S-100 for generality of metadata elements. Product specifications can restrict the specifications of the framework S-100 standard as needed.
3) For specific comments refer to the comments to S-101 Phase 4 draft. since the metadata clauses in the latest version of S-101 (Phase 4) adhere closely to S-100 metadata.
4) A metadata register has been discussed before by TSMAD and should be pursued.
5) Clarify that product specifications can also provide for different metadata for different delivery modes (transfer set, service, message, etc.) as appropriate.

Part 5 (Feature Catalogue)

(See the “GFM Revisions” section earlier in this document.)

FC builder needs to be made available and tested,

This Part needs to be updated to conform with the changes discussed in early 2011 (about TSMAD22).

Part 7 (Spatial Schema)

Some domains use other primitives like circle, this might occur even in marine protected areas, which could possibly be notified as a radius about a location.

Example 1: Some U.S. marine protected areas are defined as the area within a 20 NM radius seaward of a specified point.

Example 2: Area notices (AIS Area notice – (Binary / Application Specific) Message 8 FI 22 defines one type of area as a circle, given by center (latitude/longitude), radius, and scale factor (by which the dimensions are multiplied); or a rectangle in terms of southwest corner coordinates, E and N dimensions, scale factor, and orientation (rotation in the clockwise direction).

Recommendation

1) Other primitives defined in ISO 19107 should be added, perhaps as an additional conformance class. (Domain experts with a need should develop proposals.)

Part 10 (Encoding Formats)

Provisions for other Encodings

A provision for including “format definition files” in an exchange set is needed. This is intended to cover XML schemas and validation rules (for example, Schematron rules as described in ISO 19757-3). Format definition files like XML schemas will be another type of file in the exchange set, or can they be just other kinds of support files?

(These changes would presumably be part of the definition of XML/Gml encoding so presumably actual actions on this would be recommended in that extension.) (The primary question is whether delivery of such schemas is in fact needed as part of the transfer set. because data would be validated at the production stage and the feature catalogue would be available to validate ad-hoc data.)

Recommendations

1) Address the issue and state the outcome of this discussion in S-100.
2) S-100 implicitly allows inline and referenced spatial, this should be formalized.

Part 11 (Product Specifications)

Data Packaging and Delivery

Packaging and delivery can be in other forms than exchange sets or data sets. For example:

Data streams: E.g., IVEF …

Messages: E.g., AIS …

Web service: Maritime safety information and navigational warnings on agency web sites…

It may be necessary to split this section into product specs, data specs and service specs product is say ENC, data more like route, service would be IVEF. Part 10 could then also reflect this differentiation which could link to conformance classes.

In addition to data interchange by transfer, ISO 19109 also defines data interchange by transactions. S-100 is focused on data interchange by transfer. Depending on how much detail S-100 chooses to provide data packaging and delivery could well become a new Part altogether.

(What changes to data packaging and delivery will be needed for the new portrayal model? E.g., additional files, file naming conventions, etc…)

New: Levels of compliance or conformance classes

S-100 should define "core" and "additional conformance class(es), including with some of the proposed extensions. The additional conformance classes should be used to define conformance with (1) spatial schema and coordinate reference systems (2) encoding formats – ISO 8211, GML (perhaps at different levels, e.g., basic geometry to more complex types), non-GML XML, (others?) (3) modelling of non-geospatial information using other models than the object-attribute paradigm.

Conformance class	Spatial	Encoding formats	Support for ????
1	Geometry level 3a HCRS EPSG4326 only	S-100 GML	Basic graphic file formats – TIF Basic text formats (TXT, HTML, CSS)
2	Geometry level 3a Any HCRS	ISO 8211, S-100 GML	All the above + Additional graphic file formats (JPG, GIF) Additional text formats (XML, XSD)
3	Geometry level 4 Any HCRS	ISO 8211, S-100 GML, others	All the above + others specified in a PS

Recommendation: Use this as starting point of discussion.

Issues not addressed in this paper

This paper does not address the following, because TSMAD has either addressed them in earlier sessions or is working on solutions:

- Portrayal
- Modeling of dates and times (a separate paper on the temporal model is being prepared).
- Revisions to Part 4c – Metadata – Data Quality (left to DQWG).
- Feature register: A separate paper is being prepared.

Action Required of TSMAD

TSMAD is invited to:

- designate sub-divisions of S-100 (e.g, parts, sub-parts, or clauses) as “core” and other conformance levels
- consider the changes required to reflect service centric data exchange within S-100 in addition to its current product centric approach (exchange sets)
- review the recommendations made in this paper for each Part of S-100 and incorporate those agreed by TSMAD into the next extension or correction of S-100, as appropriate.
- designate a work item leader for S-100 Edition 2.0.0.