Paper for Consideration by TSM

PRIMAR Catalogue and Dataset Versioning (summary for TSM discussion)

Submitted by:	PRIMAR
Executive Summary:	A mechanism for linking a dataset to a version of a catalogue (Portrayal Catalogue, Feature Catalogue) is currently missing in S-100. The issue was discussed at S100WG8 and forwarded to TSM10 for further discussion. This paper explores further the outcome of a GitHub discussion concerning mentioned documents, and how the issue may be solved by introducing commonly agreed upon principles for Product Specification elements versioning.
Related Documents:	S-101PT11_2023_11.1_EN_S-100_Versioning_Discussion_V1 S100WG8-24 Catalogue and dataset versioning S-100 Github: https://github.com/iho-ohi/S-100WG/issues/10#issuecomment-1957729426
Related Projects:	

Introduction / Background

The PRIMAR paper to S-100WG 8 highlighted the inability to use Dataset Discovery Metadata attribution to relate a dataset to a specific FC/PC version. The suggested extension of DatasetDiscoveryMetadata to solve the problem was not accepted, and the S-100WG discussion quickly focused on the versioning of and between the elements* of a Product Specification.

* Data Product Specification (DPS), Data Classification and Encoding Guide (DCEG), Feature Catalogue (FC), Portrayal Catalogue (PC) and Validation.

The following action was captured from S-100WG8:
In principle, the meeting agreed to a consistent versioning approach for product specification and their catalogues. The versioning between a product specifications main document, feature catalogue and portrayal catalogue should be consistent (equal version numbering). However, the meeting recognized the need to further pursue this issue seeking a final decision on the matter on the upcoming TSM meeting. [Action X/X]: Provide further discussion material to coming TSM meeting.

This paper explores further the outcome of a GitHub discussion on the topic, ref related documents. It describes how the issue may be solved by introducing commonly agreed upon principles for Product Specification elements versioning.

Basis for the S-100WG discussion was also a paper for discussion at S-101PT11 in 2023 which further elaborated on these elements.

Of relevance is also what S-100 offers guidance on versioning. S-100 versioning (clarification, revision, new edition; and their denotation) as laid out by IHO TR 2/2007 and adopted by S-100 is further described in Annex A.

Analysis/Discussion
1. The S101PT paper contained the following table:

Ref	Type of change	Example	DPS	DCEG	FC	PC	Validation
1	Major change includes an S-100 version change	New concept in S-100 used by S-101	X.0.0	X.0.0	X.0.0	X.0.0	X.0.0
2	New content no portrayal impact	Attribute value added[1]	X.X.X	X.X.X	X.X.X	X.X.X	Possibly If yes: X.X.X
3	New content with portrayal impact	New attribute value and corresponding symbol	X.X.X	X.X.X	X.X.X	X.X.X	Possibly. If yes: X.X.X
4	Symbol change only data not affected	New symbol for existing content	X.X.X	X.X.X	X.X.X	X.X.X	X.X.X
5	Validation change only	Additional check non critical	X.X.X	X.X.X	X.X.X	X.X.X	X.X.X
6	Encoding change	Change of guidance to cover new real world concept consistently	X.X.X	X.X.X	X.X.X	X.X.X	Possibly. If yes: X.X.X
7	Feature catalogue correction, no content or portrayal impact	Add listed value to FC that is described in the DCEG but missing from FC.	X.X.X	X.X.X	X.X.X	X.X.X	X.X.X
8	Portrayal catalogue correction, no content or FC impact	Correct portrayal catalogue rule	X.X.X	X.X.X	X.X.X	X.X.X	X.X.X

2. The table content led to further discussion in GitHub, resulting in a revision of the table. The revision seemed to be supported at least by some of the involved people, supporting the viewpoint that **the first two numbers of the version should always be synchronous:**

Ref.	Type of change	Example	DPS	DCEG	FC	PC	Validation
1	Major change includes an S-100 version change	New concept in S-100 used by S-101	✗.0.0	✗.0.0	✗.0.0	✗.0.0	✗.0.0
2, 3	New content	Attribute value / Feature added or removed,	A.✗.0	A.✗.0	A.✗.0	A.✗.0	A.✗.0
4, 8	PC change or correction	Add a symbol, correct a rule, change a color	A.B.-	A.B.-	A.B.-	A.B.✗	A.B.-
5	Validation change (possibly resulting from ref. 6)	Additional check non critical	A.B.-	A.B.-	A.B.-	A.B.-	A.B.✗
6	Encoding change	Change of guidance to cover new real world concept consistently, editorial change	A.B.-	A.B.✗	A.B.-	A.B.-	A.B.-
7	FC correction, no content or portrayal impact	Correct attribute which is not evaluated by PC	A.B.-	A.B.-	A.B.✗	A.B.-	A.B.-
9	DPS correction/clarifications	Editorial change	A.B.✗	A.B.-	A.B.-	A.B.-	A.B.-

3. The synchronous anticipation is also supported by annex 4 of the S-100 Roadmap, namely the Dual Fuel Concept for S-100 ECDIS, where in the description of S-100 ECDIS operation in chapter 2.6.2 it is stated that: "... *The following point(s) follow as a direct consequence of the machine readable nature of such files (FC/PC authors note) in the S-100 ECDIS: 3.* **Feature and Portrayal Catalogues are constant for a single revision of a product specification**..."

4. The PRIMAR paper to S-100WG 8 highlighted the inability to use Dataset Discovery Metadata attribution to relate a dataset to a specific FC/PC version. However, if the synchronicity between DPS and FC/PC versioning is

determined to be that the two first numbers of the version always must be equal – this may be the solution solving our problem. We would then be able to anticipate that:

- For example a 2.0 version of S-101 would always be using the 2.0 version of FC/PC. For DPS revision uptick in version number (2.1,2.2 and so on) we would always expect there to be 2.1 and 2.2 versions of FC/PC available for provision.
- For example a 2.1 version of S-101 would always be using the 2.1 version of FC/PC.
- For a clarification uptick in version number in FC/PC (2.1.1,2.1.2 and so on) we would always provide the latest available version (adhering to the revision number), as the clarifications will always be backwards compatible.

5. If this is the common understanding, the relationship between a dataset and PC/FC could be anticipated and understood by the encoding in the DatasetDiscoveryMetadata:

S100_ProductSpecification

Role Name	Name	Description	Mult	Type	Remarks
Class	S100_ProductSpecification	The Product Specification contains the information needed to build the specified product	-	-	-
Attribute	name	The name of the Product Specification used to create the datasets	0..1	CharacterString	The name in the GI Registry should be used for this field. For example, "Electronic Navigational Chart"
Attribute	version	The version number of the Product Specification	0..1	CharacterString	TR 2/2007 specifies versioning of Product Specifications.
Attribute	date	The version date of the Product Specification	0..1	Date	
Attribute	productIdentifier	Machine readable unique identifier of a product type	1	CharacterString (Restricted to Product ID values from the IHO Product Specification Register, in the IHO Geospatial Information Registry)	For example, "S-101"

Example 1: Dataset S-101 2.0.
name: S-101
version: 2.0
productIdentifier: S-101
Anticipation: To be used together with FC/PC 2.0. – always using the latest clarification (e.g. 2.0.5).

Example 2: Dataset S-101 2.1.
name: S-101
version: 2.1
productIdentifier: S-101
Anticipation: To be used together with FC/PC 2.1. – always using the latest clarification (e.g. 2.1.5).

With this paper we have provided a potential solution for PRIMAR as a service provider to enable the use of DatasetDiscoveryMetadata in combination with anticipated logic for Product Specification elements versioning, to instantly understand which versions of FC/PC that should be used together with a dataset.
This could enable a service provider to facilitate Catalogue distribution alongside the dataset distribution.

6. In order to make the anticipated logic work, a couple of principles for the versioning of Product Specification elements need to be established:
   A. The first two numbers of the version should always be synchronous.
   B. For Catalogue (FC/PC) Clarifications (2.1.1,2.1.2 and so on) it is always the latest available version that should be expected used together with a dataset, as the clarifications will always be backwards compatible.

If the principles are agreed upon, they should probably be further described and added, either in S-100 Part 17 or in S-98.

7. The link from a dataset to a version of the Interoperability Catalogue should probably also be part of the consideration, but this is not further elaborated as the Interoperability Catalogue is not part of a Product Specification.

8. An additional observation.
The challenge of S100_ProductSpecification attribute Version having multiplicity [0..1] has been identified. It should be mandatory, but a change in multiplicity cannot be introduced before edition 6 of S-100. A change proposal will be provided at a later stage. The workaround if Version information is not provided would be to extract the information from the dataset encoding (DSID/PRED subfield in the 8211 encoding, productIdentifier field in the GML encoding and productSpecification (embedded metadata in root group) in HDF5.

Conclusions
- The logical principles of versioning elements of a Product Specification should be agreed upon:
  o The first two numbers of the version should always be synchronous.
  o For Catalogue (FC/PC) Clarifications (2.1.1,2.1.2 and so on) it is always the latest available version that should be expected used together with a dataset, as the clarifications will always be backwards compatible.
- The anticipated logic should be further described and added, either in S-100 Part 17 or in S-98.

Action Required of TSM
The TSM is invited to:
- Note the paper and discuss the proposed principles concerning the versioning of elements of a Product Specification.
- Consider including logical principles in S-100 Part 17 or in S-98.
- Take any action appropriate.

Annex A

The versioning at S-100 level is defined in S-100 Part 12 and implements the versioning definitions from IHO Technical Resolution TR 2/2007. The concept of Clarification, Revision and New Edition is introduced, with specific requirements for version control.

Clarifications are non-substantive changes to S-100. Typically, clarifications: remove ambiguity; correct grammatical and spelling errors; amend or update cross references; insert improved graphics in spelling, punctuation and grammar. A clarification must not cause any substantive semantic change to S-100. Clarifications are the responsibility of the relevant subordinate body and may be delegated to the responsible editor.
Clarifications shall be denoted as n.n.n. Each clarification or set of clarifications approved at a single point in time shall increment n by 1.

Revisions are defined as substantive semantic changes to S-100. Typically, revisions change existing specifications to correct factual errors; introduce necessary changes that have become evident as a result of practical experience or changing circumstances; or add new specifications within an existing section. Revisions could have an impact on either existing users or future users of a revised standard. It follows that a full consultative process that provides an opportunity for input from as many stakeholders as possible is required. Proposed changes to S-100 should be evaluated and tested wherever practicable. The approval of Member States is required before any revisions to S-100 can enter into force. All cumulative clarifications must be included with the release of approved corrections revisions. A revision shall not be classified as a clarification in order to bypass the appropriate consultation processes.
Revisions shall be denoted as n.n.0. Each revision or set of revisions approved at a single point in time shall increment n by 1. Revision version control shall set clarification version control to 0.

New Editions of S-100 introduce significant changes. New Editions enable new concepts, such as the ability to support new functions or applications, or the introduction of new constructs or data types, to be introduced. New Editions are likely to have a significant impact on either existing users or future users of the revised standard. It follows that a full consultative process that provides an opportunity for input from as many stakeholders as possible is required. Proposed changes to S-100 should be evaluated and tested wherever practicable. The approval of Member States is required before any New Edition of S-100 can enter into force. All cumulative clarifications and revisions must be included with the release of an approved New Edition of S-100.
New Editions shall be denoted as n.0.0. Each new edition approved at a single point in time shall increment n by 1. New Edition version control shall set the clarification and revision version control to 0.

Although Part 12 specifically states that "… This Part excludes the maintenance regime of Product Specifications", the maintenance and versioning concepts defined here have been adopted by most Product Specifications produced to date.

In addition to Part 12, S-100 also points to IHO TR 2/2007 for versioning of Product Specifications:

S100_ProductSpecification

Role Name	Name	Description	Mult	Type	Remarks
Class	S100_ProductSpecification	The Product Specification contains the information needed to build the specified product	-	-	-
Attribute	name	The name of the Product Specification used to create the datasets	0..1	CharacterString	The name in the GI Registry should be used for this field. For example, "Electronic Navigational Chart"
Attribute	version	The version number of the Product Specification	0..1	CharacterString	TR 2/2007 specifies versioning of Product Specifications.
Attribute	date	The version date of the Product Specification	0..1	Date	
Attribute	productIdentifier	Machine readable unique identifier of a product type	1	CharacterString (Restricted to Product ID values from the IHO Product Specification Register, in the IHO Geospatial Information Registry)	For example, "S-101"

The versioning guidance as laid out in TR 2/2007 (PRINCIPLES AND PROCEDURES FOR MAKING CHANGES TO IHO TECHNICAL STANDARDS AND SPECIFICATIONS), and adopted into S-100 as described above, does not describe:

1. Any versioning rules between the different elements* of a Product Specification.

* Data Product Specification (DPS), Data Classification and Encoding Guide (DCEG), Feature Catalogue (FC), Portrayal Catalogue (PC) and Validation.