

INTERNATIONAL HYDROGRAPHIC ORGANIZATION



IHO GUIDELINE STANDARD FOR CREATING S-100 PRODUCT SPECIFICATIONS PART A

Version 0.1

2018-01-31

Special Publication No. S-???

Guideline for Creating An S-100 Product Specification

Part A - Content

Published by the
International Hydrographic Organization
MONACO

Contents

PART A - CONTENT.....	1
1. OVERVIEW.....	1
1.1. INTRODUCTION.....	1
1.2. REFERENCES.....	1
1.3. TERMS AND ABBREVIATIONS.....	1
2. S-100 PRODUCT SPECIFICATION TEMPLATE AND ITS COMPONENTS.....	1
2.1. GENERAL S-100 CONCEPTS IMPORTANT TO THE READABILITY OF THE PRODUCT SPECIFICATION.....	2
2.1.1. MANDATORY VERSUS OPTIONAL REQUIREMENTS	2
2.1.2. CAMELCASE AND ITS USE IN S-100	2
2.1.3. WHAT ARE MULTIPLICITIES AND HOW ARE THEY USED IN S-100	3
2.2. MAIN PARTS TO A S-100 PRODUCT SPECIFICATION.....	3
2.2.1. THE OVERVIEW SECTION AND ITS SUB-ELEMENTS.....	3
2.2.1.1. CREATION OF THE DATA PRODUCT SPECIFICATION.....	3
2.2.1.2. TERMS AND DEFINITIONS.....	3
2.2.1.3. ABBREVIATIONS	4
2.2.1.4. ACRONYMS.....	4
2.2.1.5. INFORMAL DATA PRODUCT DESCRIPTION	4
2.2.2. USE OF LANGUAGE	4
2.2.3. PRODUCT SPECIFICATION MAINTENANCE	4
2.2.3.1. NEW EDITION	4
2.2.3.2. REVISIONS.....	4
2.2.3.3. CLARIFICATION	5
2.2.4. VERSION NUMBERS.....	5
2.2.5. SPECIFICATION SCOPES	5
2.2.6. DATASET IDENTIFICATION	6
2.2.7. DATA CONTENT AND STRUCTURE	7
2.2.7.1. FEATURE BASED DATA CONTENT STRUCTURE.....	7
2.2.7.2. COVERAGE BASED DATA CONTENT STRUCTURE	8
2.2.8. DATA PRODUCT FORMAT.....	8
2.2.8.1. 8211	8
2.2.8.2. GML.....	9
2.2.8.3. HDF5	10
2.2.8.4. OTHER ENCODINGS AND ENCODING PROFILES	11
2.2.9. APPLICATION SCHEMA.....	11
2.2.9.1. FEATURE CATALOGUE.....	12
2.2.9.2. MODEL ELEMENTS USED IN A PRODUCT SPECIFICATION	12
2.2.10. REFERENCE SYSTEMS / CRS.....	14
2.2.11. DATA QUALITY	14
2.2.11.1. DATA QUALITY IN THE DATA.....	15
2.2.11.2. DATA QUALITY CHECKS	16
2.2.12. DATA CAPTURE INSTRUCTIONS	16
2.2.13. MAINTENANCE	17
2.2.13.1. INCREMENTAL UPDATE.....	17
2.2.13.2. WHOLE DATASET UPDATE	18

- 2.2.13.3. SUPPORT FILE MAINTENANCE18
- 2.2.14. DATA PRODUCT DELIVERY18
- 2.2.15. METADATA.....19
- 2.2.15.1. DISCOVERY METADATA20
- 2.2.15.1.1. DISCOVERY METADATA FOR DATASETS.....21
- 2.2.15.1.2. DISCOVERY METADATA FOR SUPPORT FILES.....21
- 2.2.16. PORTRAYAL21
- 2.2.17. ADDITIONAL INFORMATION22
- 3. IHO S-100 INFRASTRUCTURE22
- 3.1. GI REGISTRY.....22
- 3.2. FEATURE CATALOGUE BUILDER.....23
- 3.3. DCEG BUILDER.....24
- 3.4. PORTRAYAL CATALOGUE BUILDER25

Part A - Content

1. Overview

Creating an S-100 based product specification can be a big challenge for groups with little experience with S-100. Especially since S-100 is a comprehensive framework with great detail that may need to be considered. Therefore, having a guide to assist development teams through the process can help significantly and decrease the time it takes to learn S-100 and the specifics required when creating an S-100 based product specification.

1.1. Introduction

This guideline is intended to serve as a guide for anyone planning to develop an S-100 compliant product specification. The guideline consists of two main parts; Part A (this document) is an in-depth description of the various components of an S-100-based product specification, and Part B (a separate document) describes the typical steps and activities involved in creating an S-100-based product specification.

1.2. References

IHO S-58 ENC Validation Checks, Edition 6.0.0, May 2017

IHO S-99 Operational Procedures for the Organization and Management of the S-100 Geospatial Information Registry, Edition 1.1.0, November 2012

IHO S-100 Information Paper, January 2011

IHO S-100 Universal Hydrographic Data Model Editions 3.0.0 and 4.0.0 (in preparation)

IHO S-122 Marine Protected Areas, 2017 (draft)

IHO S-123 Maritime Radio Services, 2017

1.3. Terms and Abbreviations

GML	Geography Markup Language
IALA	International Association of Lighthouse Authorities
IEC	International Electrotechnical Commission
IHO	International Hydrographic Organization
IMO	International Maritime Organization
RENC	Regional ENC Coordinating Centre
XML	eXtensible Markup Language
XSD	XML Schema Definition

2. S-100 Product Specification Template and Its Components

A data product specification is a precise technical description which defines a geospatial data product. It describes all the features, attributes and relationships of a given application and their mapping to a dataset. It includes general information for data identification as well as information for data content and structure, reference system, data quality aspects, data capture, maintenance, delivery and

metadata. It may be created and used on different occasions, by different parties and for different reasons.

Part 11 of S-100 describes data product specifications for geographic data products. Its aim is to provide a clear and similar structure for any data product specification to be written. A product specification shall constitute a set of human readable documentation. Generally, it should also include machine readable files for information such as the feature catalogue, the application schema and the CRS parameters. An example of a compliant product specification is shown in Appendix 11-B of S-100. In addition to a 'human readable' document, it is possible to create a machine readable (e.g. XML) summary of the Product Specification, which can be an XML document of the tables found in Part 11 of S-100.

2.1. General S-100 Concepts Important to The Readability of The Product Specification

2.1.1. Mandatory Versus Optional Requirements

For a product specification to claim compliance with S-100, some specific parts are required. For example, it is required to include a feature catalogue, while a portrayal catalogue is optional. Tables in part 11 indicate what is mandatory and what is optional by the multiplicity column in each table. Below in an example of such a table.

Name	Description	Multi	Type
title	Official designation of the data product	1	CharacterString
abstract	Informal description of the data product	1	CharacterString
acronym	Any acronyms for the title of the data product	0..*	CharacterString
content	Textual description of the content of any dataset which conform to the specification	1	CharacterString
spatialExtent	Description of the spatial extent covered by the data product	1	EX_Extent (ISO 19115)
temporalExtent	Description of the temporal extent covered by the data product	0..1	EX_Extent (ISO 19115)
specificPurpose	Specific purpose for which the data shall be or has been collected	1	CharacterString

Table 2-1 Example of S-100 table (Informal Description of the Data Product, S-100 Table 11-1)

In the above example, the Multi column indicates which elements are mandatory and which are optional. The rows, title, abstract, content, spatial extent and specific purpose are all mandatory, as indicated by the multiplicity of 1, while the rows acronym and temporal extent are optional, as indicated by the multiplicities of 0..* and 0..1.

2.1.2. CamelCase and Its Use In S-100

S-100 uses camelCase extensively. Camel case (stylized as camelCase or CamelCase; also known as camel caps or more formally as medial capitals) is the practice of writing compound words or phrases such that each word or abbreviation in the middle of the phrase begins with a capital letter, with no intervening spaces or punctuation. Common examples include "iPhone ", "eBay", "FedEx", "DreamWorks", "HarperCollins", "iCarly", "WordWorld", and "WordGirl". It is also sometimes used in online usernames

such as "JohnSmith", and to make multi-word domain names more legible, for example in advertisements. [Wikipedia, 2017]. S-100 makes use of camelCase as a method to construct distinct identifiers or names of elements used within S-100 itself, within the GI Registry, Feature Catalogues, etc.

The camelCaseIdentifier must:

- Be compound words in which the words are joined without spaces and are capitalized within the compound.
- Be unique within the registry.
- Conform to ISO 646 with uppercase characters A-Z, 0-9, " _ ", and lowercase characters a-z.
- Features and Information types must begin with uppercase A-Z.
- Attributes and enumerated values must begin with lowercase a-z.

Example 1 BeaconCardinal is the Camel Case identifier for the feature Beacon Cardinal

Example 2 categoryOfLandmark is the Camel Case identifier for the attribute Category of Landmark

2.1.3. What are Multiplicities and How Are They Used in S-100

Within the tables, models and other parts of S-100, the concept of multiplicity is used to give the range of a particular entity in the given context. Such as how many times shall an attribute be used within a class. More details can be found in S-100 Part 1-4.5.3.3 S100_Multiplicity and subsequent paragraphs.

2.2. Main Parts to a S-100 Product Specification

This section gives a highlight of the parts that make up an S-100 product specification, and elaborates on why these parts are needed.

2.2.1. The Overview Section and its Sub-elements.

The overview section of a product specification provides a reader with general introductory information about the data product together with product specification metadata. S-100 states that the Overview shall include the following parts:

2.2.1.1. Creation of the Data Product Specification

information about the creation of the data product specification, which shall include the title, a reference date, the responsible party and the language of the document. Additionally, information about the maintenance regime for the product specification should also be included. This can be a statement about a regular review of the specification, or that it will be updated on a as needed bases, etc.

2.2.1.2. Terms and Definitions

A section of terms and definitions used within the product specification. These are often a useful reference for the reader, and should reflect the content of the specification as well as the context it is tended to be used in.

2.2.1.3. Abbreviations

Any abbreviations used in the specification should be listed with their full meaning in a separate section within the introductory parts of the product specification.

2.2.1.4. Acronyms

Any acronyms for the name of the data product, for example AML Additional Military Layer, or ENC Electronic Navigational Chart.

2.2.1.5. Informal Data Product Description

An informal description of the data product, which can read like an abstract of the specification, it's purpose and intended use context. See also 2.2.6.

2.2.2. Use of Language

Although optional, it can be beneficial to add a Use of language section to elaborate the intended meaning of specific words used within the product specification document, including appendices and annexes. The purpose is to remove as much ambiguity as possible regarding these words so that the specification is clear regarding what is a requirement, what is highly recommend and what is optional. In this regard, the following statements have been used frequently within IHO specifications.

- “Must” indicates a mandatory requirement.
- “Should” indicates an optional requirement, that is the recommended process to be followed, but is not mandatory.
- “May” means “allowed to” or “could possibly”, and is not mandatory.

2.2.3. Product Specification Maintenance

Changes to a product specification issued by IHO will be released as a new edition, a revision, or as a document that includes clarification. It is likely that other issuing authorities will follow IHO's example. Generally, this text follows the guidance given in S-100 Part 12, where three types of change are described; new edition, revision and clarification. These change types are elaborated in a general manner below. Minor variations to these can be excepted depending on the type of product specification.

2.2.3.1. New Edition

New Editions 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. New Editions are likely to have a significant impact on either existing users or future users of a product specification.

2.2.3.2. Revisions

Revisions are defined as substantive semantic changes. Typically, revisions will introduce change to correct factual errors; introduce necessary changes that have become evident as a result of practical experience or changing circumstances. A revision must not be classified as a clarification. Revisions could

have an impact on either existing users or future users of this specification. All cumulative clarifications will be included with the release of approved corrections revisions.

Changes in a revision are minor and ensure backward compatibility with the previous versions within the same Edition. Newer revisions, for example, introduce new features and attributes. Within the same Edition, a dataset of one version could always be processed with a later version of the feature and portrayal catalogues. In most cases a new feature or portrayal catalogue will result in a revision of this specification.

2.2.3.3. Clarification

Clarifications are non-substantive changes. Typically, clarifications: remove ambiguity; correct grammatical, spelling errors, and punctuation; amend or update cross references; and insert improved graphics. Clarification must not cause any substantive semantic changes.

Changes in a clarification are minor and ensure backward compatibility with the previous versions within the same Edition. Within the same Edition, a dataset of one clarification version could always be processed with a later version of the feature and portrayal catalogues, and a portrayal catalogue can always rely on earlier versions of the feature catalogues.

Changes in a clarification are minor and ensure backward compatibility with the previous versions.

2.2.4. Version Numbers

The associated version control numbering to identify changes (n) in S-100 and derived specifications generally follow this format:

New Editions denoted as n.0.0

Revisions denoted as n.n.0

Clarifications denoted as n.n.n

The same format for versioning is also being adopted in several of IHO's other standards.

2.2.5. Specification Scopes

Some parts of a product specification may apply to the whole product whereas other parts of the product specification may apply to parts of the product. For example, a coordinate reference system will generally apply to the complete product, while maintenance regimes may be different for navigational features versus contextual features. This difference would be described using scopes which state what parts of a specification are global and which parts are specific to a portion of the specification. When this is the case for a product being specified, the specification scopes section defines the various "scopes" within the overall product specification, and how they should be identified in the datasets. Depending on the type of data product specification, the scope may include items in in the table below. (Note that some of the elements in this table are from ISO 19115:2006 and have been renamed in ISO 19115-1:2014, and editions of S-100 after Edition 3.0.0 may use the new elements.)

Name	Description	Mult	Type
scopeIdentification	Specific identification of the scope	1	CharacterString
level	Hierarchical level of the data specified by the scope	0..1	MD_ScopeCode (ISO 19115)
levelName	Name of the hierarchy level	0..1	CharacterString
levelDescription	Detailed description about the level of the data specified by the scope	0..1	CharacterString
coverage	Subtype of a feature that represents real world phenomena as a set of attributes	0..1	CharacterString
Extent	Spatial, vertical and temporal extent of the data	0..1	EX_Extent (ISO 19115)

Table 2-2 - Specification Scope Information (S-100 Table 11-3)

If a specification is homogeneous across the whole data product it is only necessary to define a general scope (root scope), to which each section of the product specification applies. This general scope may look something like the following example.

Scope ID: Global
Level: 006- series
Level name: ENC Dataset

The product specification may specify a partitioning of the data content of the product on the basis of one or more criteria. Such partitioning may be different for different parts of the data product specification. Each such part of the data content shall be described by a specification scope that may inherit or override the general scope specification. In principle, any or all of the remaining sections of the product specification may have variants which apply to the scopes within the product. Each variant must identify the scope(s) to which it applies.

EXAMPLE Data products to support navigation often contain two sets of feature types: those that provide navigation information that changes rapidly and is essential for safety of navigation, and those that provide background reference information. Maintenance and delivery information would be partitioned on the basis of these groupings; reference system information would not.

2.2.6. Dataset Identification

In addition to the informal description of the data product (see also 2.2.1.5), S-100 also calls for a section that describes information that uniquely identifies any dataset as being created in accordance with a specific product specification series.

Different from the general information about the data product, the dataset identification is for the individual dataset. It is possible to standardize some of the elements if that is beneficial. For example, the purpose attribute value may be common among all datasets created from a particular specification. Other attributes may benefit from following a common schema, such as the dataset title may follow a particular style that users are familiar with. Some of the attributes are codelist types defined elsewhere,

such as in ISO 19115, these attributes are limited to the values given in those codelists. There may be cases where it is beneficial to restrict the given codelists to a subset of values, if for example not all values make sense for the scope of the specification.

This information is stored in the metadata that is associated with the dataset. Therefore, it is important to ensure that appropriate metadata attributes are available, and to harmonize this section with the metadata section.

Some product specifications have merged the informal description of the data product with this section, into a common section. This is an option that can be considered for new specifications also.

2.2.7. Data Content and Structure

The data content and structure of products creating from an S-100 product specification is defined in an application schema. Application schemas are fundamental elements of any S-100 based product specification. The General Feature Model part of S-100 (Part 3) specifies the rules for developing an application schema which includes the conceptual model for features and their characteristics and associations.

2.2.7.1. Feature Based Data Content Structure

The data content of a geographic application is defined in accordance with a view of real world features and in the context of the requirements of a particular application. The content is structured in terms of objects. S-100 considers two types of object:

- 1) Features – features are defined together with their properties. A feature is an abstract representation of real world phenomenon. Features have two aspects; feature type and feature instance. A feature type is a class and is defined in a Feature Catalogue. A feature instance is a single occurrence of the feature type and represented as an object in a data set.
- 2) Information Types – information types are used to share information among features and other information types. An information type is a class of object which is defined in a Feature Catalogue. An instance of an information type is an identifiable unit of information in a data set. Information types have only thematic attribute properties. An instance of an information type may be associated with one or more feature instances or other instances of information type. Information types can be thought of as shared attributes.

The General Feature Model provides a conceptual model for these objects. The definitions for object types are held in a Feature Catalogue. The GFM also acts as a conceptual model for the Feature Catalogue. Spatial information is defined in S-100 Part 7, Spatial Schema, and consists of simple geometry which can be expressed in multiple configurations. The application schema must define the spatial components used in a product specification and the relationship to the feature classes.

2.2.7.2. Coverage Based Data Content Structure

Although the conventional approach is to consider an image or a grid as a unique entity on its own, and to not consider a feature structure, it is proper to consider imagery, gridded and coverage data as feature oriented data. In the simplest form, an image or any set of gridded data can be considered as a single feature. Thus, rules for application schema for feature data apply to imagery and gridded data. However, care must be taken to ensure that the application schema accurately defines the Imagery and Gridded Data Spatial Schema in accordance with S-100 Part 8 Clause 8-6 and the Gridded Data Spatial Referencing as defined in Clause 8-8. If the product contains a series or set of images or gridded data sets, then the application schema defining the spatial relationships should be defined as specified in S-100 Part 8 Clause 8-7.

2.2.8. Data Product Format

The S-100 based product specifications shall define the format (encoding) in which each scope within the data product is delivered. This includes a description of file structures and formats where applicable, or the format of a data stream if so applicable. The encoding structure could be specified completely in the specification, or by reference to a separate profile or standard. S-100 includes profiles of three encodings; ISO 8211 binary encoding, GML (ISO 19136) encoding and HDF5 encoding. One of these can be referenced by a product specification along with a description of how to use them within the specific product specification. For example, a given product would have a specific GML application schema, expressed in one or more XML Schema Definition Language files. Specialized products may use other encodings, for example S-101 contains a profile of ISO 8211 binary encoding.

A brief description of the S-100 profiled encodings is provided in the subsequent three sections.

2.2.8.1. 8211

The ISO/IEC 8211 Specification is a data descriptive file format for information interchange. S-100 Part 10a specifies the structure of an exchange set at the record and field levels. It further specifies the contents of the physical constructs required for their implementation as ISO/IEC 8211 data records, fields, and subfields. The grouping of records into ISO/IEC 8211 files is considered application specific and is, therefore, necessary to be described in the relevant product specification. In S-100 only the binary ISO/IEC 8211 format is used.

Field Tag: DSID	[Upd] *	Field Name: Data Set Identification
------------------------	---------	-------------------------------------

Subfield name	Label	Format	Subfield content and specification
Record name	RCNM	b11	{10} **)
Record identification number	RCID	b14	Range: 1 to 2 ³² -2
Encoding specification	ENSP	A()	Encoding specification that defines the encoding
Encoding specification edition	ENED	A()	Edition of the encoding specification
Product identifier	PRSP	A()	Unique identifier for the data product as specified in the product specification
Product edition	PRED	A()	Edition of the product specification
Application profile	PROF	A()	Identifier that specifies a profile within the data product
Data set name	DSNM	A()	The name of the data set
Edition number	EDTN	b12	The edition number of the data set
Update number	UPDN	b12	The update number of the data set
Issue date	ISDT	A(8)	The issue date Format: YYYYMMDD according to ISO 8601

Figure 2-1 Example of Field Tables

2.2.8.2. GML

The Geography Markup Language (GML) is an XML grammar defined by the Open Geospatial Consortium (OGC)/ISO 19136 to express geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. It should be noted that the concept of feature in GML is a very general one and includes not only conventional "vector" or discrete objects, but also coverages and sensor data. The ability to integrate all forms of geographic information is key to the utility of GML.

Part 10b specifies a profile of GML meant to be used as a basis for the development of GML application schemas for S-100 data products. The GML application schema for each data product defines a file format for the machine-to-machine exchange of information structured in conformance with the application schema for the data product, as defined in the appropriate product specification.

The S-100 GML profile defines the core GML components that shall be used in GML encodings for S-100 data products. This profile defines a restricted subset of XML and GML types that excludes GML features not required by S-100 GML datasets. This subset of GML is then used to create the specific GML encoding for a product specification. This approach is described in Figure 2-2.

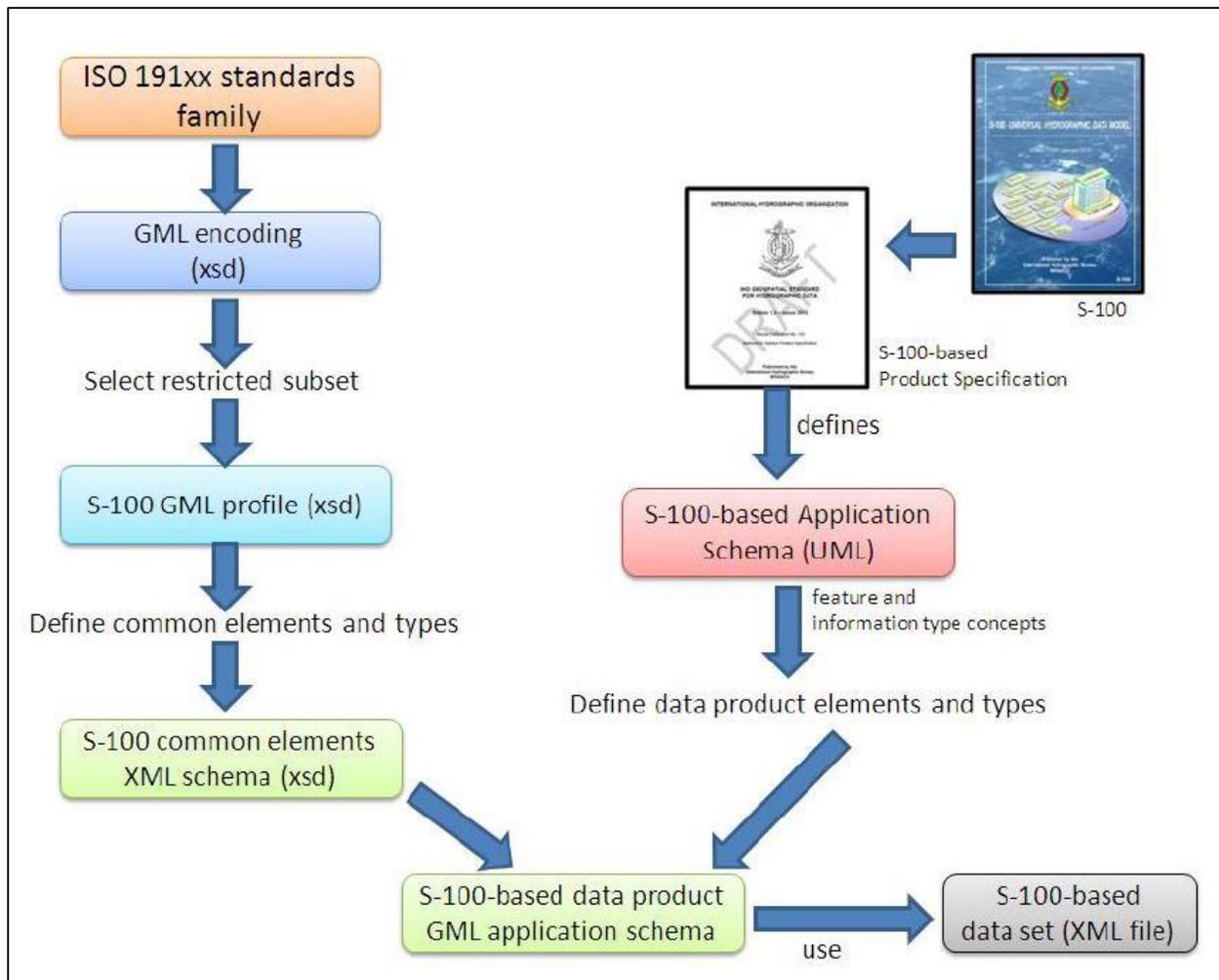


Figure 2-2 Derivation of profile and its use by a data product

2.2.8.3. HDF5

The Hierarchical Data Format 5 (HDF5) HDF has been developed by the HDFgroup as a file format for the transfer of data that is used for imagery and gridded data. HDF5 is particularly good at dealing with data where complexity and scalability are important. Data of virtually any type or size can be stored in HDF5, including complex data structures and data types.

S-100 Part-10c specifies an interchange format to facilitate the moving of files containing data records between computer systems.

The HDF5 profile in S-100 Edition 3.0.0 lacks a common structure and does not provide a mapping between GFM and the encoding. This will be provided in the S-100 Edition 4.0.0 profile of HDF5.

```
(96)
Group size = 6
Number of attributes = 36
dataCodingFormat = 2
dataProtection = 0
dateTimeOfFirstRecord = 20150827T010000Z
dateTimeOfIssue = 20160914T011100Z
dateTimeOfLastRecord = 20150827T060000Z
depthTypeIndex = 2
gridLandMaskValue = -1.0
gridOriginLatitude = 43.14162
gridOriginLongitude = -79.8209
gridSpacingLatitudinal = 0.04499
gridSpacingLongitudinal = 0.0614
horizDatumReference = ESPG
horizDatumValue = 4326
maxSurfCurrentSpeed = 1.29
methodCurrentsProduct = POMGL_3d_model
minGridPointLatitude = 0
minGridPointLongitude = 0
minSurfCurrentSpeed = 0.01
nameRegion = US_Great_Lakes
nameSubregion = Lake_Ontario
numPointsLatitude = 25
numPointsLongitude = 61
numberOfNodes = 0
numberOfStations = 0
numberOfTimes = 6
productSpecification = S-111_version_1.0.0
protectionScheme = Not_Applicable
surfaceCurrentDepth = 0.0
timeRecordInterval = 3600
typeOfCurrentData = 6
uncertaintyOfDirection = -1.0
uncertaintyOfHorzPosition = -1.0
uncertaintyOfSpeed = -1.0
uncertaintyOfTime = -1.0
uncertaintyOfVertPosition = -1.0
verticalDatum = 0
```

Figure 2-3 Display of a HDF5 sample file (using HDFView) containing the 36 dataset root directory attributes

2.2.8.4. Other Encodings and Encoding profiles

Where the S-100 profiled encodings are not sufficient for the intended use of a product specification, another encoding can be specified or profiled within the product specification. Sufficient detail should be given to permit implementers an easy understanding of the encoding chosen. A significant implication of this approach is that any system applications supporting the resulting data product may require bespoke development, which may be difficult in standardized equipment, such as ECDIS. An alternative is to request additional encoding profiles be added to S-100, which then can be added to systems that comply with S-100.

2.2.9. Application Schema

Ideally, the full application schema is described in the Data content and structure section. It must be expressed in UML (S-100 3.0.0 11-7.1), which is described in S-100 Part 1, Conceptual Schema Language. The S-100 Product Specification template stipulates that specifications that have large application schemas need contain only specific examples in UML in the specification document, since the application schema is realized in full in the feature catalogue. Because application schemas generally become too big to remain easily readable in one page, it may be beneficial for overall readability to split up application schemes into sections based on functions and elements. The elements used (e.g all feature types in one diagram and all enumerated lists in another diagram) are discussed in a later

section. Functions can be things like specific topics (e.g. how buoys are modelled, or how contact details are modelled) and how these are constructed.

2.2.9.1. Feature Catalogue

ISO 19110 defines a feature catalogue as a catalogue that contains definitions and descriptions of the feature types, feature attributes, and feature associations occurring in one or more sets of geographic data. Therefore, the feature catalogue acts as a machine-readable representation of the application schema, and gives a system the means to describe the elements of a dataset that conforms to the feature catalogue.

Because of the role as a machine-readable representation of the application schema, it can substitute in the product specification for the application schema in its entirety when the application schema may be too large to present in a document in its entirety in the form of UML diagram(s).

It is possible to create Feature Catalogues in a variation of ways, but it is recommended that it is done using the IHO Feature Catalogue Builder, as described in paragraph 3.2.

2.2.9.2. Model Elements Used in a Product Specification

This section is used for elaborating the data elements used in a product specification. This is done to help readability of the document as a stand-alone document. For example, the Product Specification Template goes into detail of what types may be used and somewhat copy the S-100 descriptions. This method is given in the example below;

Example

Feature Types

<The following clauses describe the different feature types that may be used in the feature catalogue.>

Geographic

<Geographic (geo) feature types form the principle content of the dataset and are fully defined by their associated attributes and information types.>

Meta

<Meta features contain information about other features within a data set. Information defined by meta features override the default metadata values defined by the data set descriptive records. Meta features must be used to their maximum extent to reduce meta attribution on individual features.>

Feature Relationship

<A feature relationship links instances of one feature type with instances of the same or a different feature type. There are three common types of feature relationship: Association, Aggregation and Composition >

Information Types

<Information types are identifiable pieces of information in a dataset that can be shared between other features. They have attributes but have no relationship to any geometry; information types may reference other information types.>

A second option would be to describe model elements with references to S-100, but with the consequence that the reader of the product specification need to review S-100 for details. It should be noted that this approach might complicate the readability of the document instead of what would be the case were the definitions copied or paraphrased within the product specification.

A third option utilized by some product specifications is to group elements according to some logical scheme, and then describe those groupings. This method allows a combination of the type description and at the same time link it with the usage within the specifications. Below is an example of this method, describing showing all information types in a specific product specification.

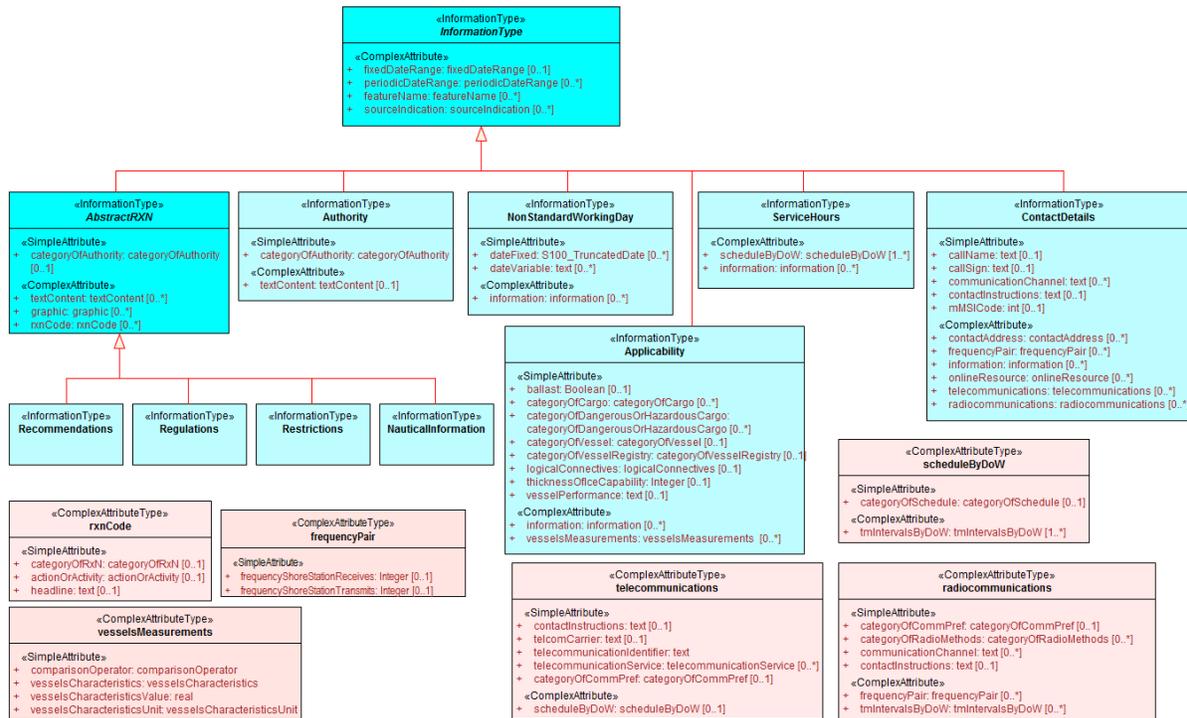


Figure 2-4 Overview of S-123 Information Types

Ultimately the choice of how to best describe the application schema and the components that it consists of rests with the product specification developers.

2.2.10. Reference systems / CRS

Product specifications must specify the reference systems that are to be used by the data products that result from the specification. All S-100 based product specifications that include georeferenced information will have a horizontal reference system, while vertical reference systems are for specifications that yield data products that include height information or bathymetry. Some specification includes more than one vertical reference system, such as one for sounding data and one for height data. Mainly two methods are available for how S-100 Product Specification should specify reference systems. These are to either specify the reference system used in full, or more commonly by referencing an already specified reference system. Part 6 of S-100 provides information of how to describe and specify a reference system. The more common method of simply referencing an already specified reference system is generally done by establishing a convention in the specification by stating the reference system, or list of references systems used, and then by adding the information in the metadata of the resulting dataset. Below is an example from the S-100 product specification template.

horizontalDatumReference	Reference to the register from which the horizontal datum value is taken	1	characterString	EPSG
horizontalDatumValue	Horizontal Datum of the entire dataset	1	Integer	4326
verticalDatum	Vertical Datum of the entire dataset	1	S100_VerticalAndSoundingDatum	

Figure 2-5 Reference Systems

As noted in the example in Figure 2-5, the EPSG registry is a useful register of horizontal datums. The codelist value for WGS84, the most commonly used horizontal datum is 4326. For vertical datums and sounding datums, S-100 include an enumerated list called S100_VerticalAndSoundingDatum. The most commonly used vertical and sounding datums are included here. The enumerated list can be extended by requesting IHO-S-100WG to include additional values to the list.

2.2.11. Data quality

All S-100 based product specifications should include comprehensive ways of capturing information about the quality of the data that results from the specification. The data quality indicators are required for users and user systems to assess fitness for use of the data. Moreover, the data products may be used by user groups that it was not originally intended for, and with good quality indicators these secondary users can make better assessments of the data's usefulness in their application.

There are several types of data quality elements that should be considered for any S-100 based product specifications, these elements include, but are not limited to;

- intended purpose of data;
- statement of quality or lineage;
- completeness of the data in terms of coverage;
- logical consistency;
- positional uncertainty and accuracy;
- temporal accuracy;

2.2.11.2. Data quality checks

It is recommended to develop conformance checks, or validation checks for each product specification. These checks serve as a list of tests to validate datasets creating to conform to the product specification. The checks should include tests for data format consistency, conformance to the specification and data model logical consistency. S-58 is the standard for validating S-57 ENC's, and has been improved over almost 20 years of experience with S-57 production and use. It can service a great template for how to structure checks and inspiration for what type of checks to consider when creating a new specification. A major lesson from S-57 and S-58 is that validation checks take time to develop, and they develop significantly with experience. It should therefore be expected that the conformance checks will evolve as experience is gained with any product specification and they should therefore be considered a regular maintenance item.

No	Check description	Check message	Check solution	Conformity to:	Apply to
100	For each feature object where its geometry is not COVERED_BY a DataCoverage	Objects fall outside the coverage object.	Ensure objects are not outside of the limits of the cell.	PS 10.9	B, S
101	If the cell file size is greater than 20 Megabytes.	The cell is larger than 20Mb in size.	Ensure that the cell is not larger than 20Mb.	PS 11.3	B, S?

Figure 2-7 Example of validation checks from draft S-122 v1.0.0

2.2.12. Data Capture Instructions

Any S-100 product specification must provide information on how the data is to be captured. This information should be as detailed and specific as necessary. To this end, the S-100 Product Specification Template recommends the development of a Data Classification and Encoding Guide, (DCEG). The DCEG is used to link practical examples to the data model. It would for example explain how different types of underwater rock is to be encoded using a specific data model, including which feature class should be used, and what attributes and their expected values that correspond to the different types of underwater rock. The DCEG template is a tabular representation of the feature catalogue with option to include graphics and remarks specific for each feature class and information class. The tabular form includes the definition of each class, the attributes that the class carry, as well as descriptions of any associations the class may have.

Expedience with ENC has shown that the data capture guide is mostly used by the data producers and serves as collective instructional document of globally common rules of how to create data. The same experience has also show that the document will grow with experience as more special cases gets resolved into a globally agreed process. This also improves overall consistency among producers and products, leading to more stable user systems as all stakeholders gain a common understanding of how to use the data products, all captured in the data capture guide.

2.2.13. Maintenance

Generally, it is expected that data created from an S-100 Product Specification will not remain valid indefinitely. It is therefore necessary to specify how data created in accordance with a specification shall be maintained, including how datasets and support files are updated.

Two main types of updating routines can be expected;

- as needed; this type of updating routine means that datasets are updated when there is a need to do so, and are to be considered the current information till there is an update. Electronic Navigational Charts and Nautical Publications are two types of data that generally is maintained in this manner.
- by schedule; this type of updating routine means that datasets are updated on a fixed schedule and that users can always anticipate when new datasets become available. Surface current and water level information are two types of data that generally is maintained in this manner.

Once it has been established how updating routines is to be done for products created from a product specification, it is necessary to establish the means by which to achieve this. Also, here there are mainly two options; incremental updates and whole dataset replacement, both of which are elaborated below.

It should be noted that these two types of updating are not mutually exclusive, and that combinations of the two can be utilized depending on the foreseen needs of the particular product specification.

However, it is important to specify exactly how the updating methodology will work, to enable machine processing of new data.

Data Sources or events that may impact dataset maintenance is useful to elaborate in a product specification so that both producers and users can better understand what may cause new data to be issued.

2.2.13.1. Incremental update

This type of updating is a method of updating a previously issued dataset by amending only a part of the dataset. This method can be very useful where bandwidth is an issue and the changes are relatively minor within the scope of the whole dataset. An example can be the addition of two objects to a dataset which contain thousands of other objects. The incremental update would then be a much smaller dataset that contains only the revision instructions to the main, or base dataset. Once the revision instructions are applied, the updated dataset would include the additional two objects. It may be required to retain both the base dataset and the incremental update within the user system for traceability purposes, and where this is the case, the specification should be written in a manner to explicitly make state this for implementers.

It should be noted that with incremental updates there may come a point when there are so many changes that it makes sense to re-issue the dataset in a manner that includes the changes received via incremental update, and thereby replacing the original base dataset with a new fully updated base dataset, and from there issue any changes as new incremental changes. In S-100 Edition 3, only ISO 8211 encoding supports this type of updating. But from version 4 of S-100 also GML will support this updating methodology.

2.2.13.2. Whole dataset update

This type of updating is a method of updating a previously issued dataset by replacing it wholesale with a new dataset. This method makes most sense when the replacement data alters all or a sufficiently large portion of the previous dataset. For example, when a forecasted data of a certain natural phenomenon is replaced with updated forecast data and the update data invalidates by being more current. All encodings in S-100 Edition 3 supports this method of updating.

2.2.13.3. Support file maintenance

Support files updating in S-100 based product specifications can be done through the metadata that goes with the support files. This is done by including issue date and management information in the discovery metadata file that follows the exchange set that includes the new or updated support file. Below is an example of how such instructions can be written for a product specification.

Example

The type of support file is indicated in the “purpose” field of the discovery metadata. Support files carrying the “deletion” flag must be removed from the system. When a feature pointing to a text, picture or application file is deleted or updated so that it no longer references the file, the system software must check to see whether any other feature referenced the same file, before that file is deleted.

To avoid complex management routines, it may be advantageous to specify that each support file should only be used once in the exchange set and to store the support files in a separate folder within the exchange set.

2.2.14. Data product delivery

The product specification may define the delivery medium (such as DVD or a web service) for each identified scope in the specification. This is an optional section, but it includes the structure of delivered data products, and is therefore of importance where data is delivered to systems that include a level of data import automation to allow implementers to program the system to a standardized delivery structure. It can also be useful to specify when data products can be delivered in different formats, such as SENC delivery. Data being delivered to ECDIS and similar systems generally expect exchange sets. S-100 includes a description of an S-100 exchange set for the interchange of geospatial data and its relevant metadata, and details can be found in Part 4a, Appendix D.

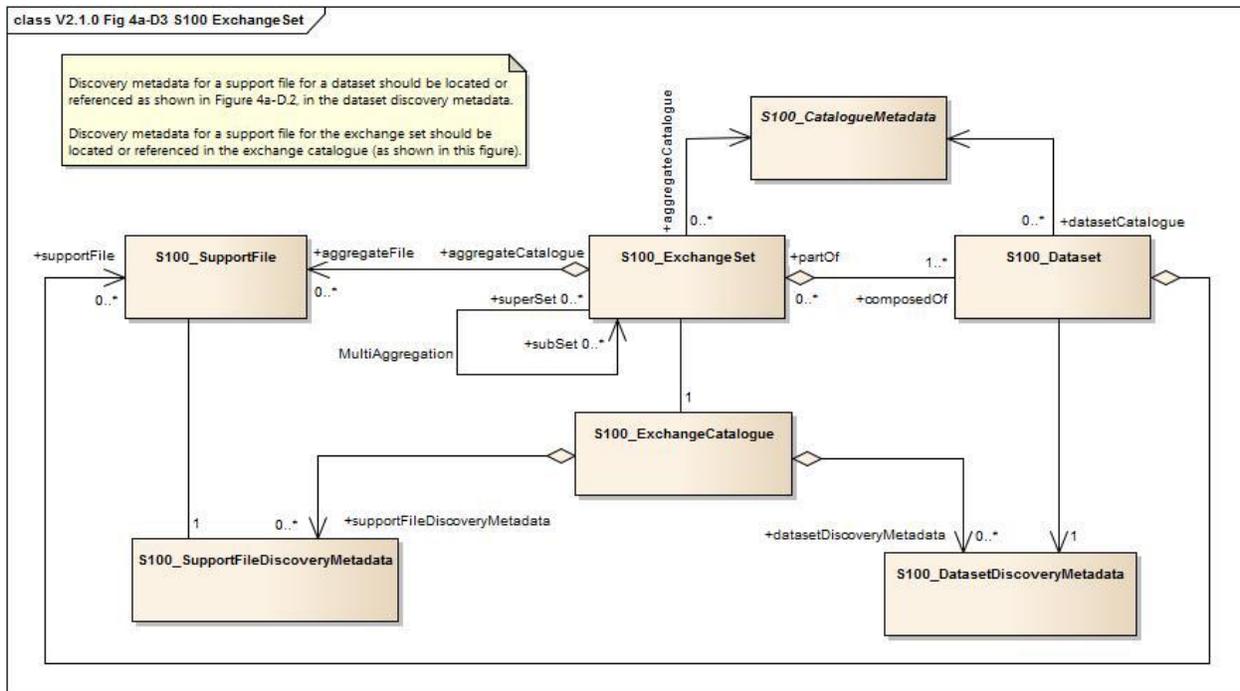


Figure 2-8 S-100 Exchange Set

Within an exchange set, several types of files are possible, in addition to the dataset. These are generally called support files. The support file types can be divided into two types; support files for the individual dataset and support files for the exchange set. Support files that are associated with the individual dataset usually includes file types such as text files and image files, while support files that are associated with the exchange set, usually are feature and portrayal catalogues.

Depending on the target user, data products may be delivered in a variety of supply chain methods, such as via Regional ENC Coordinating Centre (RENC), service providers, web service, FTP etc. It can be useful to consider the supply chain when specifying the data product delivery.

2.2.15. Metadata

Metadata is data about data. In S-100 the primary purposes of metadata are to provide information about the identification, spatial and temporal extent, quality, application schema, spatial reference system, and distribution of digital geographic data. It is applicable to the cataloguing of datasets, clearinghouse activities, and the full description of geographic and non-geographic resources.

Metadata can satisfy a number of uses; data discovery, distribution and on-line references (URL) for on-line viewing, data use, details of data creation, data fitness, data sharing, data management.

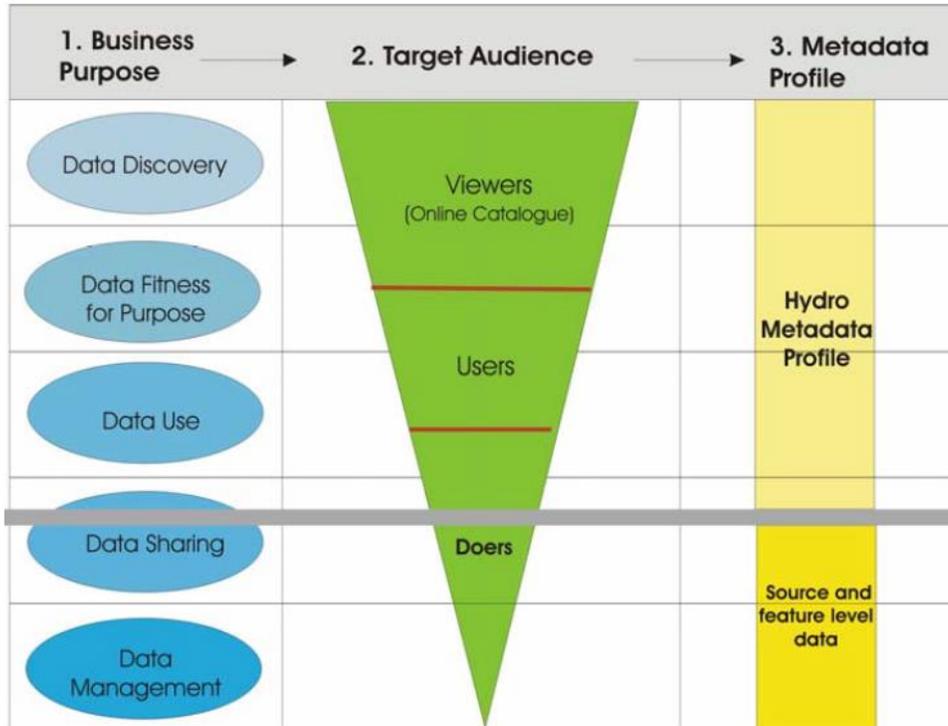


Figure 2-9 Example levels of metadata

S-100 specifies that the core metadata elements as defined in S-100 Part 4 (Metadata) shall be included with the data product. Moreover, S-100 requires that Discovery and Quality metadata shall be structured as per S-100 Parts 4a and 4c, respectively. Any additional metadata items required for a particular product specification must be documented in the data product specification, and that these should be defined using ISO 19115 and ISO 19139, with extensions or restrictions if required. The application schema shall show how metadata is carried in the datasets. This information shall be specified for each identified scope within the product specification. Furthermore, where the resulting data product is intended for a standardized user environment, like ECDIS, it should be noted that any significant deviation (e.g. addition or reduction to the standard metadata) may not be readily useable in the system depending on how the standard S-100 schemas have been implemented. Caution is therefore urged when developing the metadata for a product specification.

Specification developers should also note that newer editions of S-100 may adopt ISO 19115-1 metadata, which will result in some changes in the ISO 19115:2006 metadata used in S-100 Edition 3.0.0.

2.2.15.1. Discovery Metadata

For information exchange, there are several categories of metadata required: metadata about the overall exchange catalogue, metadata about each of the datasets contained in the catalogue, and metadata about the support files that make up the package. This is called discovery metadata in S-100

and it is used within the exchange set to enable users to learn about the content without having to open each dataset.

2.2.15.1.1. Discovery metadata for datasets

S-100 specifies that discovery metadata for datasets is contained within an external XML file created in accordance with the S-100 metadata schema. This metadata set complies with the core metadata, and extends it in a few places to provide more detail for example about reference datums and issue dates of the data. More details can be found in the normative Appendix 4a-D Discovery Metadata for Information Exchange Catalogues of S-100.

2.2.15.1.2. Discovery metadata for support files

S-100 specifies that discovery metadata for support files is contained within an external XML file created in accordance with the S-100 metadata schema. This metadata set complies with the core metadata, and extends it to provide information about the management of the support file in order to update these. More details can be found in the normative Appendix 4a-D Discovery Metadata for Information Exchange Catalogues of S-100.

2.2.16. Portrayal

Portrayal is an optional part of a product specification. However, if consistent portrayal across all user platforms is important to an S-100 based data product, then specifying how portrayal is done becomes mandatory. Within S-100 Product Specifications this is in part done by including a portrayal catalogue. The Portrayal Catalogue is a collection of defined portrayals for a feature catalogue, and includes portrayal functions, symbols, and portrayal context.

Two types of portrayal catalogues are possible in S-100. S-100 Edition 3.0.0 only specifies XSLT as the mapping language of the portrayal catalogue, however from S-100 4.0.0 an alternative approach with LUA is included for use. Part 9 of S-100 provides instructions for how a product specification can include an input Schema derived from the abstract schema provided, a set of mapping rules (defined in XSLT or LUA), a set of symbols, linestyles, colors etc. and make it available for use with product datasets. Portrayal catalogues can be created in a variety of ways, including manually and by using a Portrayal Catalogue Builder, see **Portrayal Catalogue Builder** chapter 3.4 for more details.

The IHO Hydrographic Standards and Services Committee has assigned the Nautical Cartography Working Group (NCWG) to assist with symbol design and colour selection for IHO product specifications. The NCWG chair can be contacted for further details.

Portrayal catalogues should be provided in an exchange set, and may be combined with a feature catalogue and datasets. The exact method for distribution should be defined in the product specification, but considerations should be given to efficient distribution and aims of reducing data volume wherever possible. It may therefore be beneficial to consider some form of central distribution of portrayal catalogues.

The product specification should include instructions for implementers in the use of the portrayal catalogue, including context for the use of the data. Testbeds can be a good tool in learning what types of instruction is needed for the implementers. Beyond testbeds, it can be expected that practical use will reveal additional issues that should be included in subsequent versions of a product specification.

Many of the IHO product specifications will be used in systems that have some degree of type approval requirements. Instructions for the classification society conducting the type approval should be added to product specifications where this is appropriate. These instructions should include guidance for tolerances for minor deviations and what constitute a minor deviation.

2.2.17. Additional Information

The product specification should contain all information, at a sufficient level of detail, for easy implementation by the intended stakeholders. However, there may be additional considerations that impact implementers, users, and other stakeholders. These additional considerations can be added to a section or appendix called an Implementation guide or another appropriate title. This section can be used to give context of intended use, or used to elaborate on special circumstances that impact stakeholders, and so forth.

3. IHO S-100 Infrastructure

This section describes IHO Infrastructure that has been developed to support the S-100 framework and e-Navigation.

3.1. GI Registry

A registry is the entire information system (or location) in which a collection of registers is located. A register is a collection of tables in a database containing identifiers assigned to items with descriptions of the associated items. Descriptions may consist of many types of information, including names, definitions and codes.

In the case of S-100, the IHO is hosting an on-line registry engine called S-100 Geospatial Information Registry, it can be accessed here; <http://registry.iho.int>. This registry provides the facility to access and maintain the various S-100 registers. The S-100 Geospatial Information Registry contains the following principal subordinate registers:

- Feature Concept Dictionary (FCD) register
- Portrayal register
- Metadata register
- Data Producer Code register
- Product Specifications register
- Test Bed register

Note that in version 2 of the registry, the FCD Register will be replaced by a Concept Register and a Data Dictionary Register, as shown in Figure 3-1 below.

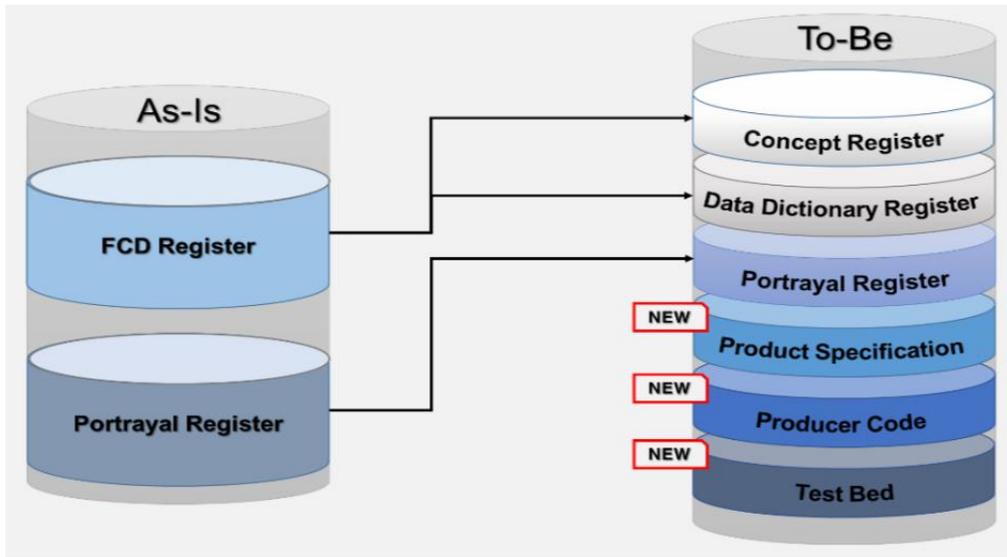


Figure 3-1 Transition from version 1 of Registry to version 2

Each register type is further sub-divided into domains e.g. hydrographic, nautical publications, ice, inland ENC etc. The administration of the registry is governed by the IHO Publication S-99 (Operational Procedures for the Organization and Management of the S-100 Geospatial Information Registry).

A major benefit of the registry and register concept is its flexibility, which allows multiple versions of similar entries in the Concept Register using unique identification and classification. An entry is classified as being either:

- valid (latest version)
- superseded (previous version/s)
- retired (no longer recommended for use)
- not valid (proposed but not accepted or no longer acceptable)

In this way Product Feature Catalogues reference items that will always be legitimate even if a newer version of the referenced item is registered at a later date. This means that if a new item is registered or an existing item is amended, newer versions of the existing product specifications are not automatically required as a consequence. The category of non-valid items is listed in the registers specifically to help identify the inappropriate reintroduction of previously rejected proposals.

3.2. Feature Catalogue Builder

A feature catalogue is a machine-readable expression of the data model for a product specification. They can either be constructed with off-the-shelf XML editors or by a Feature Catalogue Builder (FCB). Either case must comply with the structure of S-100 Part-5 and the S-100 Feature Catalogue Schema. There is a FCB available from IHO for anyone wishing to utilize it with in the creation of a feature catalogue for an S-100 product specification. The software interacts with the IHO Registry and provides

a mechanism for binding elements available in the registry together to form features and attributes, enumerated lists with their available values and so forth.

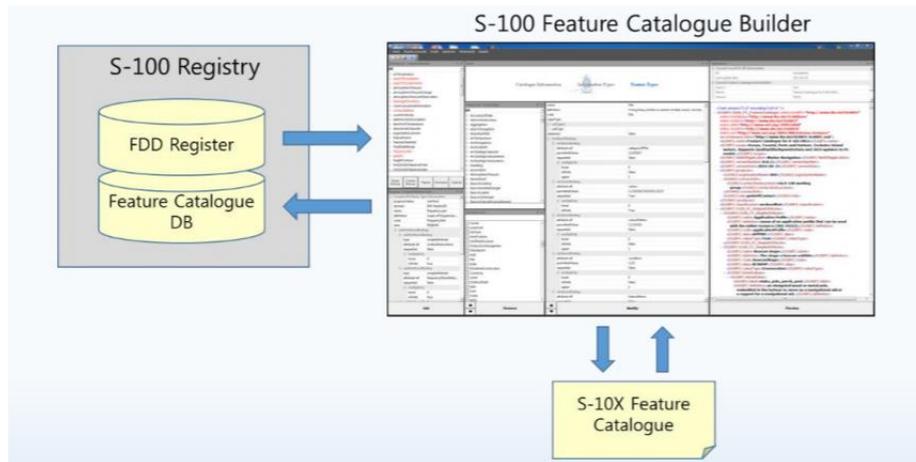


Figure 3-2 S-100 Feature Catalogue Builder

3.3. DCEG builder

To simplify the creating of the DCEG, a DCEG Builder has been created. This tool utilizes the feature catalogue to create the bindings and inputs for the DCEG tables. It is then a manual process to add images and specific text to the encoding part of the DCEG tables. Figure 3-3 below shows a high-level overview of the process to create a DCEG via the DCEG builder.

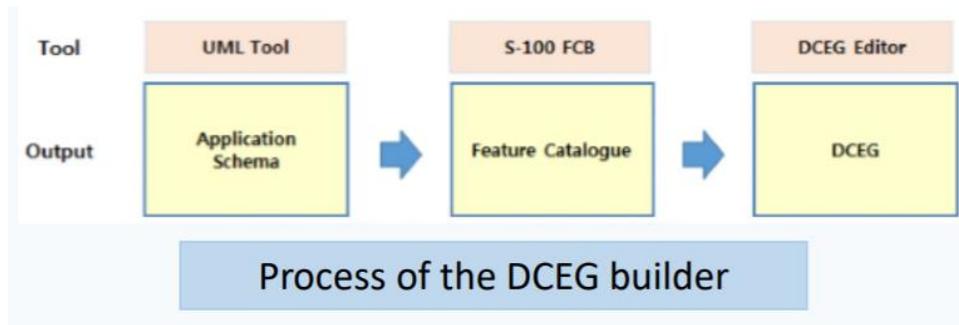


Figure 3-3 Process of creating a DCEG in the DCEG Builder

The DCEG Builder is an optional step in the creating of the product specification. Any S-100 product specification require that sufficient instructions for how to create data is included in the specification, but this information does not need to follow the DCEG style. However, the DCEG style is simple to understand and by utilizing the DCEG builder, it is also easy to create and maintain against a feature catalogue. Other methods would have to maintain the link to the feature catalogue manually, and it should be noted that this may be a significant task requiring much effort to maintain and keep current to any changes in the Feature Catalogue. Figure 3-4 below shows the location of the DCEG builder within the S-100 Infrastructure Process, and shows how the DCEG builder is integrated in the process of

creating a feature catalogue. Using the builder can greatly simplify the development process and increase consistency with the Feature Catalogue.

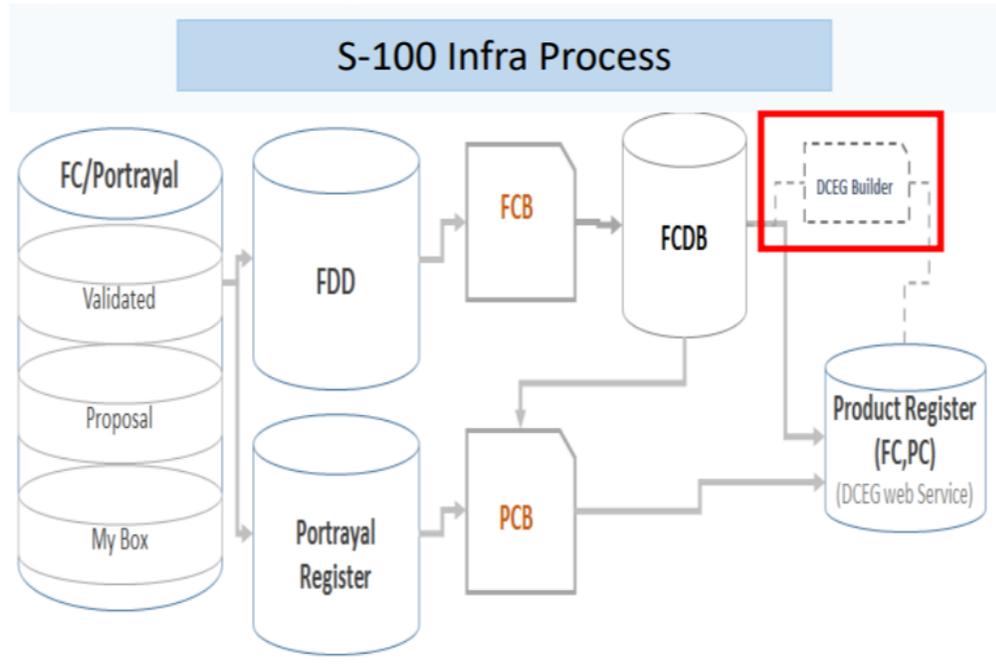


Figure 3-4 Location of DCEG builder in the S-100 Infrastructure

3.4. Portrayal Catalogue Builder

Portrayal catalogues are machine-readable instructions for how to portray data compliant with a specific data model for a specific version of a product specification. They can either be constructed manually or by a Portrayal Catalogue Builder (PCB). Either case must comply with the structure of S-100 Part-9 and the S-100 Portrayal Catalogue Schema. The IHO infrastructure includes PCB for both XSLT and LUA Portrayal Catalogues for any product specification development team wishing to utilize them with in the creation of a portrayal catalogue for an S-100 product specification. The software interacts with Portrayal Register in the IHO Registry and the Feature Catalogue, and provides an interface for binding elements available in the registry together to form symbols, line styles and area patterns for the desired elements in the feature catalogue. An overview of the interactions can be seen in Figure 3-4.