

# IHO Guidelines for Creating S-100 Product Specifications

Edition 1.1.0 – June 2020

**IHO**



International  
Hydrographic  
Organization

Published by the  
International Hydrographic Organization  
4b quai Antoine 1<sup>er</sup>  
Principauté de Monaco  
Tel: (377) 93.10.81.00  
Fax: (377) 93.10.81.40  
info@iho.int  
www.iho.int

© Copyright International Hydrographic Organization 2020

This work is copyright. Apart from any use permitted in accordance with the Berne Convention for the Protection of Literary and Artistic Works (1886), and except in the circumstances described below, no part may be translated, reproduced by any process, adapted, communicated or commercially exploited without prior written permission from the International Hydrographic Organization (IHO). Copyright in some of the material in this publication may be owned by another party and permission for the translation and/or reproduction of that material must be obtained from the owner.

This document or partial material from this document may be translated, reproduced or distributed for general information, on no more than a cost recovery basis. Copies may not be sold or distributed for profit or gain without prior written agreement of the IHO Secretariat and any other copyright holders.

In the event that this document or partial material from this document is reproduced, translated or distributed under the terms described above, the following statements are to be included:

*“Material from IHO publication [reference to extract: Title, Edition] is reproduced with the permission of the IHO Secretariat (Permission No ...../...) acting for the International Hydrographic Organization (IHO), which does not accept responsibility for the correctness of the material as reproduced; in case of doubt, the IHO’s authentic text shall prevail. The incorporation of material sourced from IHO shall not be construed as constituting an endorsement by IHO of this product.”*

*“This [document/publication] is a translation of IHO [document/publication] [name]. The IHO has not checked this translation and therefore takes no responsibility for its accuracy. In case of doubt, the source version of [name] in [language] should be consulted.”*

The IHO Logo or other identifiers shall not be used in any derived product without prior written permission from the IHO Secretariat.

### Document History

Changes to this Specification are coordinated by the IHO S-100 Working Group. New editions will be made available via the IHO website. Maintenance of the Specification shall conform to IHO Resolution 2/2007 (as amended).

| Version Number | Date        | Author  | Purpose  |
|----------------|-------------|---------|--|
| 0.1            | 31 Jan 2018 | EM, RM  | First draft  |
| 0.2            | 31 Aug 2018 | EM, RM  | Comments from NIPWG and S-100WG; multiple updates to conform to S-100 Edition 4.0.0; miscellaneous editorial updates                       |
| 1.0.0 RC1      | 11 Mar 2019 | RM; JW  | IHO Secretariat editorial changes and styling; S-100 WG4 decisions; references updated following publication of S-100 4.0.0, S-122, S-123. |
| 1.0.0          | May 2019    | S-100WG | First Edition.   |
| 1.1.0          | June 2020   | S-100WG | Substantive editorial amendments for readability. First Edition published on IHO website.  |
|                |             |         |  |
|                |             |         |  |
|                |             |         |  |
|                |             |         |  |
|                |             |         |  |
|                |             |         |  |
|                |             |         |  |

Page intentionally left blank

## **Contents**

|  |    |
|--|----|
| PART A - CONTENT.....  | 1  |
| A-1 OVERVIEW .....   | 1  |
| A-2 INTRODUCTION.....  | 1  |
| A-3 REFERENCES .....   | 1  |
| A-4 TERMS AND ABBREVIATIONS.....   | 2  |
| A-4.1 TERMS.....   | 2  |
| A-4.2 ABBREVIATIONS .....  | 5  |
| A-5 S-100 READINESS LEVELS.....  | 6  |
| A-6 S-100 PRODUCT SPECIFICATION TEMPLATE AND ITS COMPONENTS.....                             | 8  |
| A-6.1 GENERAL S-100 CONCEPTS IMPORTANT TO THE READABILITY OF THE PRODUCT SPECIFICATION ..... | 9  |
| A-6.1.1 MANDATORY VERSUS OPTIONAL REQUIREMENTS .....   | 9  |
| A-6.1.2 CAMEL CASE AND ITS USE IN S-100 .....  | 9  |
| A-6.2 MAIN PARTS OF AN S-100 PRODUCT SPECIFICATION .....                                     | 10 |
| A-6.2.1 THE OVERVIEW SECTION AND ITS SUB-ELEMENTS.....                                       | 10 |
| A-6.2.1.1 INTRODUCTION .....   | 10 |
| A-6.2.1.2 REFERENCES .....   | 10 |
| A-6.2.1.3 USE OF LANGUAGE.....   | 10 |
| A-6.2.1.4 TERMS AND DEFINITIONS .....  | 11 |
| A-6.2.1.5 ABBREVIATIONS.....   | 11 |
| A-6.2.1.5.1 ACRONYMS.....  | 11 |
| A-6.2.1.6 GENERAL DATA PRODUCT DESCRIPTION.....  | 11 |
| A-6.2.1.7 PRODUCT SPECIFICATION MAINTENANCE .....  | 11 |
| A-6.2.1.7.1 NEW EDITIONS.....  | 11 |
| A-6.2.1.7.2 REVISIONS.....   | 11 |
| A-6.2.1.7.3 CLARIFICATIONS .....   | 12 |
| A-6.2.2 VERSION NUMBERS.....   | 12 |
| A-6.2.3 SPECIFICATION SCOPES.....  | 12 |
| A-6.2.4 DATASET IDENTIFICATION .....   | 13 |
| A-6.2.5 DATA CONTENT AND STRUCTURE (APPLICATION SCHEMA) .....                                | 13 |
| A-6.2.5.1 FEATURE BASED DATA CONTENT STRUCTURE.....  | 14 |
| A-6.2.5.2 COVERAGE-BASED DATA CONTENT STRUCTURE.....   | 14 |
| A-6.2.6 DATA PRODUCT FORMAT .....  | 14 |
| A-6.2.6.1 ISO 8211.....  | 15 |
| A-6.2.6.2 GML.....   | 15 |
| A-6.2.6.3 HDF5.....  | 16 |
| A-6.2.6.4 OTHER ENCODINGS AND ENCODING PROFILES.....   | 17 |
| A-6.2.7 DATA CONTENT AND STRUCTURE.....  | 17 |
| A-6.2.8 FEATURE CATALOGUE .....  | 19 |
| A-6.2.9 DATASET .....  | 19 |
| A-6.2.10 DATASET LOADING AND UNLOADING .....   | 19 |

|   |  |    |
|---|--|----|
| A-6.2.11  | GEOMETRY.....  | 20 |
| A-6.2.12  | REFERENCE SYSTEMS.....                                   | 20 |
| A-6.2.13  | OBJECT IDENTIFIERS .....                                 | 21 |
| A-6.2.14  | DATA QUALITY.....  | 21 |
| A-6.2.15  | DATA CAPTURE AND ENCODING INSTRUCTIONS.....              | 21 |
| A-6.2.16  | MAINTENANCE.....   | 22 |
| A-6.2.16.1  | INCREMENTAL UPDATE .....                                 | 22 |
| A-6.2.16.2  | WHOLE DATASET UPDATE .....                               | 22 |
| A-6.2.16.3  | SUPPORT FILE MAINTENANCE.....                            | 23 |
| A-6.2.17  | DATA PRODUCT DELIVERY.....                               | 23 |
| A-6.2.17.1  | SERVICES AND DATA STREAMS .....                          | 24 |
| A-6.2.18  | DATASET NAMING RULES .....                               | 25 |
| A-6.2.19  | METADATA.....  | 26 |
| A-6.2.19.1  | DISCOVERY METADATA .....                                 | 27 |
| A-6.2.19.1.1  | DISCOVERY METADATA FOR DATASETS.....                     | 27 |
| A-6.2.19.1.2  | DISCOVERY METADATA FOR SUPPORT FILES.....                | 27 |
| A-6.2.19.1.3  | METADATA FOR STREAMED DATA AND SERVICES.....             | 27 |
| A-6.2.20  | PORTRAYAL.....   | 27 |
| A-6.2.21  | ADDITIONAL INFORMATION .....                             | 28 |
| A-7   | IHO S-100 INFRASTRUCTURE .....                           | 28 |
| A-7.1   | IHO GEOSPATIAL INFRASTRUCTURE (GI) REGISTRY .....        | 28 |
| A-7.2   | FEATURE CATALOGUE BUILDER.....                           | 29 |
| A-7.3   | DCEG BUILDER.....  | 30 |
| A-7.4   | PORTRAYAL CATALOGUE BUILDER.....                         | 31 |
| PART B - STEPS IN THE DEVELOPMENT OF A PRODUCT SPECIFICATION..... |  | 33 |
| B-1   | OVERVIEW.....  | 33 |
| B-2   | INTRODUCTION.....  | 33 |
| B-3   | REFERENCES.....  | 33 |
| B-4   | TERMS AND ABBREVIATIONS.....                             | 34 |
| B-4.1   | TERMS.....   | 34 |
| B-4.2   | ABBREVIATIONS .....                                      | 38 |
| B-5   | OVERVIEW OF STEPS FOR DEVELOPMENT.....                   | 38 |
| B-5.1   | BASIC DEVELOPMENT PROCESS.....                           | 38 |
| B-5.2   | REVIEW CYCLE FOR S-100-BASED PRODUCT SPECIFICATIONS..... | 42 |
| B-5.3   | PROCESS FOR EXTENDING A SPECIFICATION.....               | 42 |
| B-6   | INITIATION .....   | 43 |
| B-7   | DEVELOP DATA MODEL (APPLICATION SCHEMA).....             | 45 |
| B-7.1   | INTRODUCTION .....                                       | 45 |
| B-7.2   | STEPS IN MODEL DEVELOPMENT .....                         | 45 |
| B-7.3   | RELATIONSHIP TO THE GENERAL FEATURE MODEL .....          | 47 |
| B-7.4   | RULES FOR APPLICATION SCHEMAS .....                      | 48 |
| B-7.4.1   | APPLICATION SCHEMAS FOR VECTOR DATA .....                | 48 |
| B-7.4.2   | APPLICATION SCHEMAS FOR COVERAGE DATA .....              | 49 |

|            |   |    |
|------------|---|----|
| B-7.4.3    | ADDITIONAL RULES.....   | 49 |
| B-7.5      | OTHER CONVENTIONS AND RECOMMENDATIONS .....   | 50 |
| B-7.5.1    | REUSE AND HARMONIZATION .....   | 50 |
| B-7.5.2    | FEATURES AND INFORMATION TYPES .....  | 50 |
| B-7.5.3    | SUPERCLASSES AND SUBCLASSES.....  | 51 |
| B-7.5.3.1  | SUPERCLASSES VERSUS CATEGORY ATTRIBUTE .....  | 51 |
| B-7.5.4    | ASSOCIATIONS AND ASSOCIATION CLASSES.....   | 52 |
| B-7.5.4.1  | NAVIGABILITY, SOURCE AND TARGET.....  | 52 |
| B-7.5.4.2  | ASSOCIATION CLASSES .....   | 52 |
| B-7.5.5    | ATTRIBUTES IN GENERAL.....  | 53 |
| B-7.5.6    | CODELIST AND ENUMERATION ATTRIBUTES.....  | 54 |
| B-7.5.7    | LABELS AND DEFINITIONS FOR LISTED VALUES.....   | 54 |
| B-7.5.8    | DATA TYPES.....   | 54 |
| B-7.5.9    | CODES FOR LISTED VALUES .....   | 55 |
| B-7.6      | RECOMMENDED PRACTICES .....   | 55 |
| B-7.6.1    | REVIEWS OF MODEL ELEMENTS AND STRUCTURE.....  | 55 |
| B-7.6.2    | DIAGRAM LAYOUT .....  | 55 |
| B-7.6.3    | COLOUR CODING OF MODEL ELEMENTS .....   | 55 |
| B-7.6.4    | DOCUMENTATION TABLES .....  | 55 |
| B-7.6.5    | RECOMMENDED SOFTWARE TOOLS.....   | 55 |
| B-7.6.6    | IDENTIFICATION OF MODELS.....   | 56 |
| B-8        | DATA CLASSIFICATION AND ENCODING GUIDE.....   | 56 |
| B-9        | IHO GI REGISTRY.....  | 56 |
| B-10       | FEATURE CATALOGUE .....   | 56 |
| B-11       | DATA TRANSFER MODES AND PACKAGING .....   | 57 |
| B-12       | METADATA.....   | 57 |
| B-12.1     | METADATA FOR EXCHANGE SET PRODUCTS.....   | 57 |
| B-12.1.1   | GENERIC METADATA MODEL.....   | 57 |
| B-12.1.2   | USE OF THE IHO S-100 GENERIC METADATA MODEL AND SCHEMAS IN EXCHANGE<br>CATALOGUES ..... | 59 |
| B-12.1.3   | EXTENSIONS IN ISO METADATA FILES .....  | 62 |
| B-12.1.4   | DATA PROTECTION, AUTHENTICATION AND ENCRYPTION .....                                    | 62 |
| B-12.1.4.1 | APPLICATION OF PROTECTION MEASURES.....   | 63 |
| B-12.1.4.2 | DIGITAL SIGNATURES .....  | 63 |
| B-12.1.4.3 | COMPRESSION AND ARCHIVE FORMAT .....  | 63 |
| B-12.1.5   | METADATA FOR IMAGERY AND GRIDDED DATA .....   | 64 |
| B-12.1.6   | EMBEDDED METADATA .....   | 64 |
| B-12.2     | METADATA FOR SERVICES.....  | 64 |
| B-12.2.1   | GENERIC METADATA FOR SERVICES .....   | 64 |
| B-12.2.2   | USE OF THE IHO S-100 METADATA MODEL AND SCHEMAS .....                                   | 64 |
| B-12.2.3   | DATA PROTECTION, AUTHENTICATION AND ENCRYPTION .....                                    | 65 |
| B-12.2.4   | EMBEDDED METADATA .....   | 65 |
| B-13       | DEFINE DATA ENCODING FORMAT .....   | 65 |
| B-13.1     | SELECTION OF ENCODING FORMAT.....   | 65 |

|                             |  |    |
|-----------------------------|--|----|
| B-13.2                      | DATA FORMAT DEFINITION ARTEFACTS FOR GML .....             | 66 |
| B-14                        | PORTRAYAL ELEMENTS AND RULES .....                         | 67 |
| B-15                        | REGISTRATION OF PORTRAYAL ELEMENTS .....                   | 68 |
| B-16                        | PORTRAYAL CATALOGUE .....                                  | 68 |
| B-17                        | REFERENCE SYSTEMS.....                                     | 68 |
| B-18                        | DATA PRODUCT DELIVERY .....                                | 69 |
| B-18.1                      | DELIVERY CONTENT AND STRUCTURE .....                       | 69 |
| B-18.1.1                    | EXCHANGE SETS .....  | 69 |
| B-18.1.2                    | SERVICES .....   | 69 |
| B-18.2                      | DATASET UPDATES .....                                      | 70 |
| B-18.2.1                    | GENERAL CONSIDERATIONS FOR UPDATES.....                    | 70 |
| B-18.2.2                    | FORMAT-SPECIFIC UPDATE CONSIDERATIONS.....                 | 70 |
| B-18.3                      | SUPPORTING INFORMATION .....                               | 70 |
| B-19                        | VALIDATION CHECKS AND DATA QUALITY .....                   | 71 |
| B-19.1                      | VALIDATION CHECKS FOR DATASETS.....                        | 71 |
| B-19.2                      | VALIDATION CHECKS FOR PACKAGES .....                       | 72 |
| B-19.3                      | COMMON VALIDATION CHECKS.....                              | 72 |
| B-19.4                      | VALIDATION CHECKS FOR BASE VERSUS UPDATE DATASETS .....    | 72 |
| B-20                        | PREPARING FOR INTEROPERABILITY .....                       | 72 |
| B-21                        | SAMPLE DATA / TEST DATASETS.....                           | 73 |
| B-22                        | TESTING AND FEEDBACK.....                                  | 74 |
| B-23                        | WORK PROCESSES.....  | 74 |
| B-23.1                      | REGISTRATION AND REQUESTING AN S- NUMBER.....              | 74 |
| B-23.2                      | PROJECT TEAMS .....  | 74 |
| B-23.3                      | ITERATIVE REFINEMENT AS A DEVELOPMENT PROCESS .....        | 74 |
| B-23.4                      | MAINTENANCE OF PRODUCT SPECIFICATIONS .....                | 75 |
| PART C – DATA QUALITY ..... |  | 77 |
| C-1                         | OVERVIEW.....  | 77 |
| C-2                         | INTRODUCTION .....   | 77 |
| C-3                         | REFERENCES.....  | 77 |
| C-4                         | TERMS AND ABBREVIATIONS.....                               | 78 |
| C-4.1                       | TERMS.....   | 78 |
| C-4.2                       | ABBREVIATIONS .....  | 79 |
| C-5                         | OVERVIEW OF DATA QUALITY MEASURES.....                     | 80 |
| C-5.1                       | DATA QUALITY MEASURES .....                                | 81 |
| C-6                         | RECOMMENDATIONS FOR PRODUCT SPECIFICATION DEVELOPERS ..... | 82 |
| C-6.1                       | COMPLETENESS > COMMISSION / OMISSION.....                  | 82 |
| C-6.2                       | LOGICAL CONSISTENCY > CONCEPTUAL CONSISTENCY .....         | 82 |
| C-6.3                       | LOGICAL CONSISTENCY > DOMAIN CONSISTENCY.....              | 83 |
| C-6.4                       | LOGICAL CONSISTENCY > FORMAT CONSISTENCY.....              | 83 |
| C-6.5                       | LOGICAL CONSISTENCY > TOPOLOGICAL CONSISTENCY .....        | 83 |
| C-6.6                       | POSITIONAL ACCURACY .....                                  | 83 |
| C-6.7                       | THEMATIC ACCURACY .....                                    | 84 |



|        |  |    |
|--------|--|----|
| C-6.8  | TEMPORAL QUALITY .....                     | 85 |
| C-6.9  | AGGREGATION .....                          | 85 |
| C-6.10 | USABILITY .....                            | 86 |
| C-7    | DATA QUALITY MEASURES .....                | 87 |
| C-8    | MINIMUM STANDARD FOR DATA VALIDATION ..... | 89 |

Page intentionally left blank

## Part A - Content

### A-1 Overview

S-100, the Universal Hydrographic Data Model, is a hydrographic geospatial data standard that can support a wide variety of hydrographic-related digital data sources; and is fully aligned with mainstream international geospatial standards, in particular the ISO 19000 series of geographic standards. This alignment enables easier integration of hydrographic data and applications into geospatial solutions. S-100 is inherently more flexible than S-57 and makes provision for such things as the use of imagery and gridded data types, enhanced metadata and multiple encoding formats. It also provides a more flexible and dynamic maintenance regime for features, attributes and portrayal via a dedicated online Registry. S-100 provides a framework of components that enables the building of standardized Product Specifications for the modeling of hydrographic data, thus providing true interoperability between different data standards and systems.

S-97 is a Guideline intended for developers and maintainers of Product Specifications that utilize the IHO framework standard S-100 (Universal Hydrographic Data Model).

### A-2 Introduction

Developing S-100-based Product Specifications can be a challenge for those with little experience with S-100, especially since S-100 is a framework standard that covers a wide range of applications that may not be utilized in every Product Specification. S-97 was created by the International Hydrographic Organization to help Product Specification developers to better understand S-100 and to provide additional guidance on how to create and extend an S-100-based Product Specification.

A core aim of S-97 is to assist in the creation of harmonized Product Specifications that are used within the e-Navigation eco-system. The term e-Navigation eco-system is meant to encompass all Product Specifications created for use in IMO-defined e-Navigation systems, both on shore and at sea, such as ECDIS; but S-100 also has a wide range of applications that can extend beyond e-Navigation.

This guideline serves as a cookbook for anyone planning to develop or extend an S-100-compliant Product Specification and consists of three parts:

- Part A is an in-depth description of the various components of an S-100-based Product Specification;
- Part B describes the typical steps and activities involved in creating an S-100-based Product Specification. Part B describes the overall process; specific activities and tasks; and includes hints for solving specific problems while the Product Specification is being developed; and
- Part C describes the data quality measures deemed appropriate for use in S-100-based Product Specifications.

### A-3 References

- |          |  |
|----------|--|
| IHOIB    | IHO S-100 Information Brochure, May 2017.  |
| ISO 8211 | Specification for a data descriptive file for information interchange structure implementations. ISO/IEC 8211, 1994. |

|             |   |
|-------------|---|
| ISO 646     | Information technology – ISO 7-bit coded character set for information interchange. ISO/IEC 646, 1991.  |
| ISO 10646   | Information technology – Universal Coded Character Set (UCS). ISO/IEC 10646, 2017.  |
| ISO 19103   | Geographic information – Conceptual schema language. ISO 19103, 2005.   |
| ISO 19110   | Geographic Information – Methodology for feature cataloguing. ISO 19110, 2005.  |
| ISO 19115-1 | Geographic information – Metadata – Part 1 – Fundamentals. ISO 19115-1, 2014, as amended by Amendment 1, 2018.  |
| ISO 19136   | Geographic information – Geography Markup Language (GML). ISO 19136, 2007. (Also available as OGC 07-036 Geography Markup Language (GML) Encoding Standard. Open Geospatial Consortium Inc., 2007.) |
| S-57        | IHO Transfer Standard for Digital Hydrographic Data, Edition 3.1, November 2000.  |
| S-58        | IHO S-58 – ENC Validation Checks, Edition 6.0.0, May 2017.  |
| S-99        | IHO S-99 – Operational Procedures for the Organization and Management of the S-100 Geospatial Information Registry, Edition 1.1.0, November 2012.   |
| S-100       | IHO S-100 – Universal Hydrographic Data Model, Edition 4.0.0, December 2018.  |
| S-122       | IHO S-122 – Marine Protected Areas, Edition 1.0.0, January 2019.  |
| S-123       | IHO S-123 – Maritime Radio Services, Edition 1.0.0, January 2019.   |

Note: In this document, “S-100” means S-100 Edition 4.0.0 unless a different edition is explicitly identified.

## A-4 Terms and abbreviations

### A-4.1 Terms

#### abstract class

an object class which cannot be **instantiated**, or is designated in an information model as not allowed to be instantiated

NOTE: Subclasses of an abstract class may be either abstract or non-abstract.

#### aggregation

special form of association that specifies a whole-part relationship between the aggregate (whole) and a component part (see **composition**) [ISO 19103]

#### application

manipulation and processing of data in support of user requirements [ISO 19101-1:2014]

#### application schema

**conceptual schema** for data required by one or more **applications** [ISO 19101-1:2014]

#### association

semantic relationship between two or more classifiers that specifies connections among their instances [ISO 19103]

#### attribute

(1) named property of an entity [ISO/IEC 2382-17:1999]

NOTE: Describes a geometrical, topological, thematic or other characteristic of an entity.

(2) UML: feature within a classifier that describes a range of values that instances of the classifier may hold [ISO/IEC 19501:2005 (Adapted)]

(3) XML: name-value pair contained in an element [ISO 19136]

**base64**

an encoding designed to represent arbitrary sequences of octets in a form that allows the use of both upper- and lowercase letters but that need not be human readable [IETF RFC 4648 (restyled)]

**code**

representation of a label according to a specified scheme [ISO 19118:2011]

**codelist**

value domain including a code for a permissible value [ISO 19136]

**codespace**

rule or authority for a code, name, term or category [ISO 19136]

EXAMPLE: Examples of codespaces include dictionaries, authorities, codelists, etc.

**composition**

form of aggregation association with strong ownership and coincident lifetime as part of the whole [ISO 19103]

**conceptual model**

**model** that defines concepts of a universe of discourse [19101-1:2014]

**conceptual schema**

formal description of a **conceptual model** [ISO 19101-1:2014]

**data client**

an end-user receiving encrypted S-100-based data; the data client will be using a software application (for example, ECDIS) to perform many of the operations detailed within the S-100 protection scheme [S-100 Part 15 (adapted)]

EXAMPLE: An ECDIS user.

**data permit**

file containing encrypted product keys required to decrypt the licensed products, normally created specifically for a particular **data client** [S-100 Part 15 (adapted)]

**data server**

an organization producing encrypted data files or issuing **data permits** to **data clients** [S-100 Part 15, (adapted)]

**feature**

abstraction of real world phenomena [ISO 19101:2003]

NOTE: A feature may occur as a type, class, or an instance. Feature type or feature instance should be used when only one is meant. **Feature class** should be used in the context of a model or application schema.

EXAMPLE: The phenomenon named 'Eiffel Tower' may be classified with other phenomena into a feature type 'tower'.

**feature association**

**relationship** that links instances of one feature type with instances of the same or a different feature type [ISO 19110]

**feature attribute**

characteristic of a feature [ISO 19101]

NOTE: A feature attribute type has a name, a data type and a domain associated to it. A feature attribute instance has an attribute value taken from the value domain of the feature attribute type.

EXAMPLE 1: A feature attribute named 'colour' may have an attribute value "green", which belongs to the data type "text".

EXAMPLE 2: A feature attribute named 'length' may have an attribute value "82.4", which belongs to the data type "real".

**feature catalogue**

a catalogue containing definitions and descriptions of the **feature types**, **feature attributes** and **feature associations** occurring in one or more sets of geographic data [ISO 19110]

**feature class**

a class in an **application schema** or **model** that represents a **feature**

**identifier**

a linguistically independent sequence of characters capable of uniquely and permanently identifying that with which it is associated [adapted from ISO/IEC 11179-3:2003]

**information type**

an identifiable unit of information in a dataset with only thematic attribute properties [S-100 3-5.1.2 (adapted)]

EXAMPLE: An information type might be used to carry a Chart Note.

NOTE Information types can also be associated with each other. This could be done where there is further supplementary information that is relevant to the information type or where there is a need to translate the information. For example, a primary information object carrying a Chart Note may contain text in English and an associated supplementary information object may carry the same text in German.

**instantiate**

represent by a concrete instance [Merriam-Webster Online <https://www.merriam-webster.com/dictionary/instantiate>]

**interface**

named set of **operations** that characterize the behaviour of an entity [ISO 19119:2005]

**metadata**

information about a **resource** [ISO 19115-1]; data that defines and describes other data [ISO 11179-3:2013]

**model**

abstraction of some aspects of reality [ISO 19109-2015]

**operation**

specification of a transformation or query that an object may be called to execute [ISO 19119:2005]

NOTE: An operation has a name and a list of parameters.

**register**

set of files containing identifiers assigned to items with descriptions of the associated items [ISO 19135]

NOTE: Descriptions may consist of many types of information, including names, definitions and codes.

**registry**

information system on which a **register** is maintained [ISO 19135]

**relationship**

semantic connection among **model** elements [ISO 19103]

**resource**

identifiable asset or means that fulfils a requirement [ISO 19115-1]

EXAMPLES: Dataset, dataset series, service, document, initiative, software, person or organization.

**scheme administrator**

organization solely responsible for maintaining and coordinating the protection scheme specified by S-100 [S-100 Part 15 (adapted)]

**service**

distinct part of the functionality that is provided by an entity through **interfaces** [ISO 19119:2005]

**spatial object**

object used for representing a spatial characteristic of a feature [ISO 19107:2003]

**stream**

in online data exchange: a continuous sequence of fragmented data to be transported by a communication system [S-100]

**universe of discourse**

view of the real or hypothetical world that includes everything of interest [19101-1:2014]

**vocabulary**

terminological dictionary which contains designations and definitions from one or more specific subject fields [ISO 1087-1:2000]

**A-4.2 Abbreviations**

AIS Automatic Identification System

DQWG Data Quality Working Group

|             |   |
|-------------|---|
| ECDIS       | Electronic Chart Display and Information System     |
| ENC         | Electronic Navigational Chart                       |
| GML         | Geography Markup Language                           |
| GFM         | General Feature Model                               |
| GI registry | Geospatial Information registry                     |
| HDF         | Hierarchical Data Format                            |
| HTTP        | Hypertext Transfer Protocol                         |
| HTTPS       | HTTP Secure   |
| IALA        | International Association of Lighthouse Authorities |
| IEC         | International Electrotechnical Commission           |
| IHO         | International Hydrographic Organization             |
| IMO         | International Maritime Organization                 |
| ISO         | International Organization for Standardization      |
| RENC        | Regional ENC Coordinating Centre                    |
| REST        | Representational State Transfer                     |
| SENC        | System Electronic Navigational Chart                |
| SOAP        | Simple Object Access Protocol                       |
| SOS         | Sensor Observation Service                          |
| SSL         | Secure Sockets Layer                                |
| SVG         | Scalable Vector Graphics                            |
| TCP/IP      | Transmission Control Protocol/ Internet Protocol    |
| VTS         | Vessel Traffic Service                              |
| WSDL        | Web Services Description Language                   |
| WFS         | Web Feature Service                                 |
| XML         | eXtensible Markup Language                          |
| XSD         | XML Schema Definition                               |
| XSLT        | eXtensible Stylesheet Language Transformations      |

## A-5 S-100 Readiness Levels

A key issue when developing new Product Specifications within the S-100 framework is the ability to communicate to the wider community the completeness of the Specification and its readiness for operational use. This is also further complicated by the many different types of operational settings for Product Specifications under development, and not all of which require all S-100 components; thus the concept of S-100 Readiness Levels has been adopted by the Hydrographic Services and Standards Committee (HSSC).

The readiness levels concept shows a progression from an idea to regular use, and allows the IHO community to gain a clear understanding of whether the Specification is ready for endorsement and approval. This will also allow other non-IHO stakeholder organizations who are leveraging the S-100



framework to gauge when their Product Specifications meet an appropriate readiness level for transition to live operation.

The following table lists the prerequisite components to meet each S-100 Readiness Level. Note that it is required that all S-100-based Product Specifications conform to S-100 and both the Feature Catalogue and the Portrayal Catalogue must use the published S-100 infrastructure and process for creation and maintenance.

| Required Product Specification component  | Level 1<br>v1.0.0 | Level 2<br>v1-2.0.0 | Level 3<br>>v2.0.0                          | Level 4<br>>v2.0.0 | Level 5<br>>v2.0.0 |
|---|-------------------|---------------------|---|--------------------|--------------------|
| Main Document<br>(Defines the relevant parts of S-100 that are required for the Product Specification)  | X                 | X                   | X   | X                  | X                  |
| <i>A Default Encoding</i>   | X                 | X                   | X   | X                  | X                  |
| S-100 Compliant Feature Catalogue   | X<br>(draft)      | X<br>(updated)      | X<br>(final,<br>from IHO<br>GI<br>Registry) | X                  | X                  |
| <i>Data Classification and Encoding Guide</i>   | X<br>(draft)      | X                   | X<br>(final)                                | X                  | X                  |
| S-100 Compliant Portrayal Catalogue<br>NOTE: Not every Specification will need a Portrayal Catalogue – this should be determined as part of the development process and stakeholder feedback. |                   | X                   | X   | X                  | X                  |
| Data Quality Checks   |                   | X                   | X   | X                  | X                  |
| Test Data Sets  |                   | X                   | X   | X                  | X                  |
| <i>Data Validation (and test datasets)</i>  |                   | X                   | X   | X                  | X                  |
| Exchange Catalogue  |                   | X                   | X   | X                  | X                  |
| Encryption / Digital Signatures   |                   |                     | X   | X                  | X                  |
| Interoperability  |                   |                     | X*<br>(draft)                               | X*<br>(tested)     | X*                 |
| Alerts and Indications  |                   |                     |   | X*                 | X*                 |
| Operational data  |                   |                     |   |                    | X                  |

(X\* = ECDIS only)

Table A-5-1 – S-100 Readiness Levels

Level 1: Contains the minimum amount of components needed to commence the development of test datasets and system prototypes. This should be considered the final stage of development before demonstration begins, and would typically be Edition 1.0.0 of a Product Specification.

Level 2: Includes additional items such as data quality checks and test data sets so that the Product Specification can be demonstrated in prototype environments. This would typically map to Edition 1.n.n - 2.0.0 of a Product Specification. Depending on the end-user requirements of the Product Specification, Level 2 can be implemented in an operational context. Subsequent S-100 Readiness Levels are then dependent on operational requirements of the product within navigation systems.

Level 3: Builds on Level 2 and includes a fully featured and documented exchange catalogue and (optionally) an encryption layer for the data and implementing system. At this level, prototype systems, products or processes should be demonstrated in a real-world environment.

Level 4: Intended only for use in vessel navigation systems such as ECS and ECDIS. At this level, the developer of the Product Specification needs to ensure that documented considerations have been given to interoperability via S-98 and alerts and indications functionality. At this level, there should be a baselined and compliant system, process or product that is shown to operate or function as expected.

Level 5: System, process or product is deployed and used routinely. At this stage, data and compliant systems are readily available for operational use. This stage includes functionality for a machine-based check of up-to-datedness (that is, automatic warnings and Update Status Reports). The functionality required for up-to-datedness could be provided within individual Product Specifications or through the S-128 Catalogue of Nautical Products.

## **A-6 S-100 Product Specification template and its components**

A data Product Specification is a precise technical description that 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 the following:

- data content and structure;
- reference system;
- data quality aspects;
- data capture;
- portrayal;
- maintenance;
- delivery; and
- metadata.

Part 11 of S-100 describes how a Product Specification for geographic data products utilizing S-100 should be formatted. The aim of Part 11 is to provide a clear and consistent structure for all S-100-based Product Specifications. A Product Specification has several components – typically a main document that provides the human-readable documentation; and machine-readable files for information such as the Feature Catalogue and Portrayal Catalogue. An example of a compliant Product Specification is shown in Appendix 11-B of S-100.

## A-6.1 General S-100 concepts important to the readability of the product specification

### A-6.1.1 Mandatory versus optional requirements

In order to be compliant to S-100, Product Specifications must contain specific mandatory items. For example, inclusion of a Feature Catalogue is mandatory, while a Portrayal Catalogue is optional. S-100 utilizes the multiplicity field to aid in determining which elements are mandatory within each product specification. Table A-6.1 below is an example of such a table.

| Name            | Description  | Mult | 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 19115-1) |
| temporalExtent  | Description of the temporal extent covered by the data product                       | 0..1 | EX_Extent (ISO 19115-1) |
| specificPurpose | Specific purpose for which the data shall be or has been collected                   | 1    | CharacterString         |

Table A-6-1 – Example of S-100 table (Informal Description of the Data Product, S-100 Table 11-1)

In the above example, the Mult (short for “Multiplicity”) column indicates which elements are mandatory and which are optional. S-100 uses the following notation to indicate types of multiplicity:

- 1 – Mandatory
- 1..\* – At least one or many
- 0..1 – Either optional or only one
- 0..\* – Either optional or many

### A-6.1.2 CamelCase and its use in S-100

S-100 uses camelCase extensively, and is based on principles from ISO 19103. Camel case (stylized as camelCase or CamelCase) 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". S-100 makes use of camelCase as a method to construct distinct identifiers or names of elements used within S-100 and the IHO Geospatial Information (GI) Registry.

According to S-100 clause 2a-4.2.3, the camelCaseIdentifier must meet the following specifications:

- Be compound words in which the words are joined without spaces and are capitalized within the compound;
- Be unique within the Registry; and
- Conform to ISO 10646<sup>1</sup> with uppercase characters A-Z, 0-9, " \_ ", and lowercase characters a-z.

In addition:

- Features and Information types must begin with uppercase A-Z; and

<sup>1</sup> S-100 2a-4.2.3, specify ISO/IEC 646 (ASCII), while elsewhere 10646 (UTF-8) is used. Since it can be problematic to mix UTF-8 and ASCII, this guidance recommends to utilize only UTF-8.

- Complex and Simple Attributes, Codelists, Enumerated and Codelist 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

## **A-6.2 Main parts of an S-100 Product Specification**

This section provides a highlight of the parts that make up an S-100 Product Specification.

### **A-6.2.1 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 (except “References” and “Use of Language” but this guidance recommends inclusion of these sections too):

- Introduction
- References
- Terms, definitions and abbreviations
  - Use of Language
  - Terms and Definitions
  - Abbreviations
- General data product description
- Data Product Specification metadata
- Product Specification maintenance

#### **A-6.2.1.1 Introduction**

This section gives information about the creation of the Product Specification, which includes the subject matter and intent of the Product Specification.

#### **A-6.2.1.2 References**

This section should list standards which define integral elements of the Product Specification; or on which implementations depend, such as normative ISO or other standards. Other standards or documents containing useful information which is not an integral part of the Product Specification may be listed as informative references.

#### **A-6.2.1.3 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 documentation, including appendices and annexes. The purpose is to remove as much ambiguity as possible so that the Specification is clear regarding what is a mandatory requirement, what is highly recommended and what is optional. Within S-100, the following has been adopted for use of language:

- “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.

#### **A-6.2.1.4 Terms and definitions**

Terms and definitions are useful references that reflect the content of the Specification as well as the context for its intended use.

#### **A-6.2.1.5 Abbreviations**

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

##### **A-6.2.1.5.1 Acronyms**

It is customary to give an acronym for the name of the data product; for example AML (Additional Military Layer) or ENC (Electronic Navigational Chart). Acronyms may also be used throughout the Specification for a variety of reasons; these should be collected in an acronym paragraph at the beginning of the document to serve as a quick reference for the reader.

#### **A-6.2.1.6 General data product description**

This section is an informal description of the data product, which can read like an abstract of the Specification, its purpose and intended use context. See also clause A-6.2.4.

#### **A-6.2.1.7 Product specification maintenance**

Changes to a Product Specification issued by the IHO will be released as a New Edition, a Revision, or as a document that includes Clarifications. Which level is used depends on the nature of the change. It is likely that other issuing authorities will follow IHOs 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 changes are defined in the following sub-clauses.

##### **A-6.2.1.7.1 New Editions**

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.

##### **A-6.2.1.7.2 Revisions**

Revisions are defined as substantive semantic changes. Typically, Revisions will introduce changes to correct factual errors or introduce necessary changes that have become evident as a result of practical experience or changing circumstances. A Revision must not also be classified as a Clarification. Revisions could have an impact on either existing users or future users of the Specification. All accumulated Clarifications will be included with the release of approved Revisions.

Changes in a Revision 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 the Specification.

### A-6.2.1.7.3 Clarifications

Clarifications are non-substantive changes. Typically, Clarifications remove ambiguity; correct grammatical and spelling errors; amend or update cross references; and/or insert improved graphics, spelling, punctuation and grammar. Clarifications 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 Catalogue.

### A-6.2.2 Version numbers

The associated version control numbering to identify changes (n) in S-100 and derived Product 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 has also being adopted for most of IHOs other Standards.

### A-6.2.3 Specification scopes

Specification scopes indicate the applicability of designated portions of the Product Specification to either the entire product or 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 specification scopes. Depending on the type of Product Specification, the scope may include items in Table A-6-2.

| Name                 | Description  | Mult | Type                       |
|----------------------|--|------|----------------------------|
| scopelidentification | Specific identification of the scope   | 1    | CharacterString            |
| Level                | Hierarchical level of the data specified by the scope                            | 0..1 | MD_ScopeCode (ISO 19115-1) |
| 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 19115-1)    |

Table A-6-2 - 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 identification: | Global scope                               |
| Level:                | 006- series                                |
| Level name:           | Dataset                                    |
| Level description:    | Level applies globally to all ENC datasets |
| Coverage:             | All features in the ENC Feature Catalogue  |

Extent:                   EX\_GeographicBoundingBox  
                              westBoundLongitude: -180  
                              eastBoundLongitude: 180  
                              southBoundLatitude: -90  
                              northBoundLatitude: 90

The Level attribute is a codelist found in ISO 19115-1 called MD\_ScopeCode comprising the major components of a Specification. The Extent attribute is a class that can be any combination of the following: description string; a geographical extent (like in the example above); vertical extent; or temporal extent.

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 Product Specification. Each such part of the data content is 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 that 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 the presence of which 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; essential information would be maintained and delivered whereas reference system information would not.

#### **A-6.2.4 Dataset identification**

In addition to the informal description of the data product (see also clause A-6.2.1.6), 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. For example, the value contained in the Purpose attribute may be common among all datasets created from a particular Specification. Other attributes may benefit from following a common model or schema, while other attributes are codelist types that are defined elsewhere, such as in ISO 19115-1. These attributes are limited to the values given in those codelists and there may be cases where it is beneficial to restrict the given codelists to a subset of values.

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 the dataset identification section, into a common section. This is an allowed option.

#### **A-6.2.5 Data content and structure (Application Schema)**

An Application Schema defines the data content and structure of products under S-100. Application Schemas are expressed using the Unified Modelling Language (UML) as described in S-100 Part 1 (Conceptual Schema Language) and allow developers to implement S-100 Product Specifications in a consistent and maintainable way. The General Feature Model 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.

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 the Application Schema into sections based on functions and elements.

#### **A-6.2.5.1 Feature based data content structure**

The data content of a geographic application is defined with a view of real world features and the requirements of a particular application. The content is structured in terms of objects. S-100 considers two types of objects or features, which are defined in S-100 Part 3, clause 5.1. They are defined as follows:

- 1) Features – Features are defined together with their properties. Features are abstract representations of real world phenomena. The word ‘feature’ can be used in two senses: feature type and feature instance. A feature type is a class and is defined in the Feature Catalogue. A feature instance is a single occurrence of the feature type and is represented as an object in a data set.
- 2) Information – Information types are used to share information among features and other information types. An information type is a class of object that is defined in the 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 one or more instances of other information types. Information types can be thought of as shared attributes.

The General Feature Model (GFM) provides a conceptual model for these objects. The definitions for feature types are held in the 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.

#### **A-6.2.5.2 Coverage-based data content structure**

S-100 also defines 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, Application Schema rules for feature data also 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.

#### **A-6.2.6 Data product format**

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. Currently, S-100 includes profiles of three encodings: ISO 8211 binary encoding; GML (ISO 19136) encoding; and HDF5 encoding. A Product Specification may reference these profiles 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 by specifying the whole encoding within the Product Specification (or by referencing an established external standard, or an appropriate combination of the two). It should be noted that in such cases, implementation costs may be higher than for systems using the standard S-100 encodings.

A brief description of the S-100 profiled encodings is provided in the following clauses.

#### A-6.2.6.1 ISO 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, therefore, must be described in the relevant Product Specification. Figure A-6-1 below depicts an example of such a description. In S-100 only the binary ISO/IEC 8211 format is used.

| 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<br>Format: YYYYMMDD according to ISO 8601                         |

Figure A-6-1 – Example of Field Tables

#### A-6.2.6.2 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 the key to the utility of GML.

S-100 Part 10b specifies a profile of GML that is used as a basis for the development of GML application schemas for S-100-based 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 are used in GML encodings for S-100-based 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 A-6-2 below.

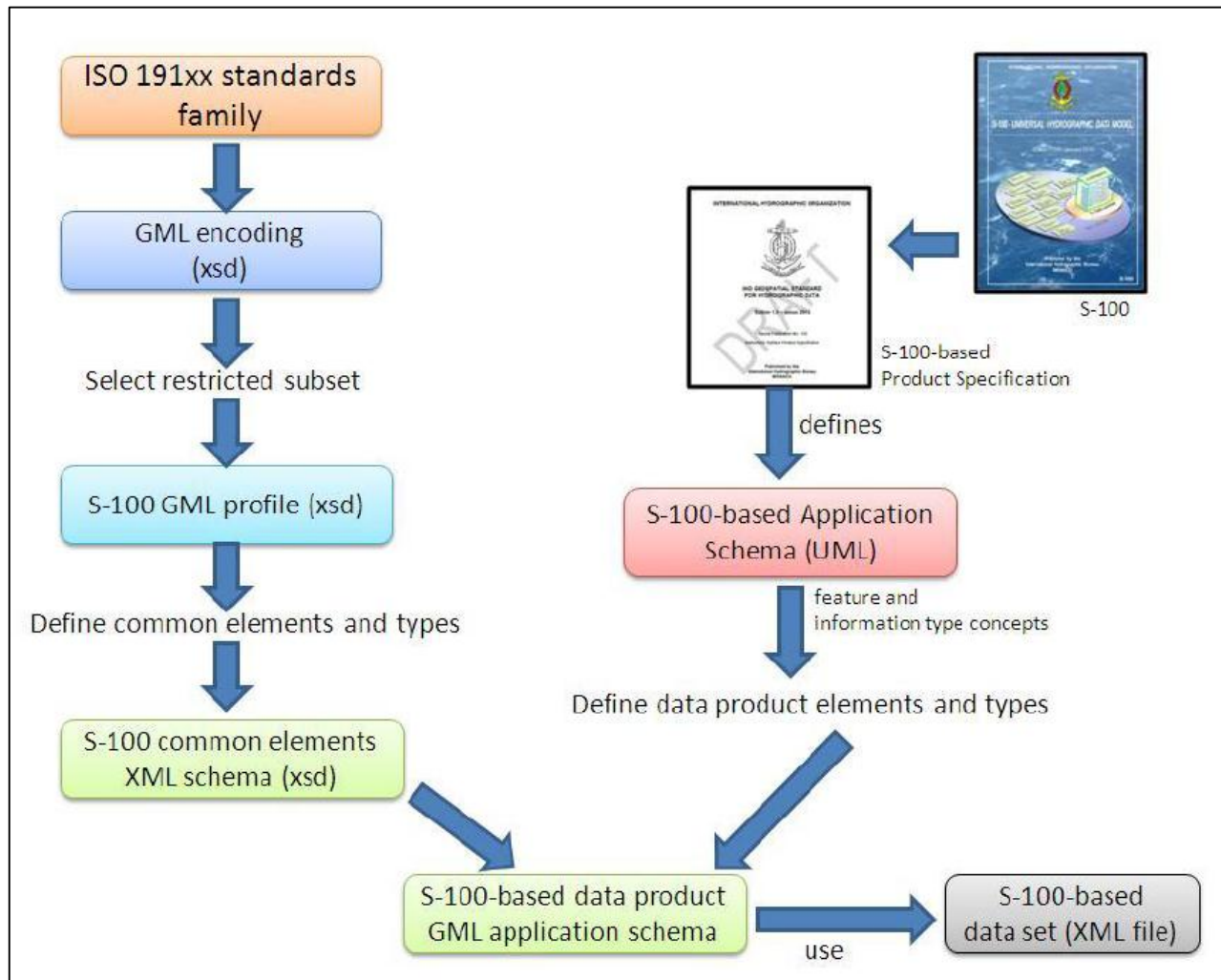


Figure A-6-2 – Derivation of the GML profile and its use by a data product

### A-6.2.6.3 HDF5

The Hierarchical Data Format 5 (HDF5) HDF has been developed by The HDF Group 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. Figure A-6-3 below depicts the structure of a typical HDF5 file.

S-100 Part 10c specifies a profile of HDF5 that is adopted for S-100. It specifies how to use HDF5 in a way that is compliant with the GFM and how to consistently specify the data formats for the types of coverages and point sets supported by S-100.

|    | 0            |               | 1            |               | 2            |               | 3            |               | 4            |               | 5            |               |
|----|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|
|    | surfaceCu... | surfaceCur... | surfaceCu... | surfaceCur... | surfaceCu... | surfaceCur... | surfaceCu... | surfaceCur... | surfaceCu... | surfaceCur... | surfaceCu... | surfaceCur... |
| 0  | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 1  | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | 225.7        | 0.05          |
| 2  | -1.0         | -1.0          | -1.0         | -1.0          | 122.5        | 0.07          | 106.0        | 0.13          | 93.7         | 0.02          | 330.2        | 0.22          |
| 3  | -1.0         | -1.0          | 160.4        | 0.06          | 8.8          | 0.1           | 348.3        | 0.07          | 324.5        | 0.2           | 322.8        | 0.34          |
| 4  | -1.0         | -1.0          | 160.9        | 0.12          | 7.1          | 0.28          | 346.1        | 0.26          | 335.3        | 0.38          | 331.3        | 0.38          |
| 5  | -1.0         | -1.0          | 217.4        | 0.1           | 309.4        | 0.24          | 346.0        | 0.3           | 350.3        | 0.37          | 338.6        | 0.44          |
| 6  | -1.0         | -1.0          | -1.0         | -1.0          | 229.6        | 0.08          | 327.8        | 0.17          | 11.3         | 0.34          | 20.0         | 0.34          |
| 7  | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | 228.3        | 0.09          | 313.7        | 0.1           | 18.5         | 0.31          |
| 8  | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | 176.3        | 0.13          | 0.8          | 0.22          |
| 9  | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | 209.4        | 0.18          | 284.4        | 0.23          |
| 10 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | 223.8        | 0.14          |
| 11 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 12 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 13 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 14 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 15 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 16 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 17 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 18 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 19 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 20 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 21 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 22 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |
| 23 | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          | -1.0         | -1.0          |

Figure A-6-3 – Image showing (left side) the structure of the file and (right side) the two-dimensional compound array of values for regularly gridded data

#### A-6.2.6.4 Other encodings and encoding profiles

When the encodings in S-100 as profiled are not sufficient for the intended use of a Product Specification, a different encoding may be specified within the Product Specification itself. If an encoding that is not part of S-100 is used, then sufficient detail should be given to permit implementers an easy understanding of the chosen encoding.

NOTE: If a non S-100-based encoding is chosen, then systems that support S-100 may not be able to handle data products that utilize this encoding (for example, ECDIS). An alternative is to request additional encoding profiles be added to S-100, which can then be added to systems that comply with S-100.

If a non S-100 encoding is used, then the metadata must be correctly encoded to indicate that the Product Specification contains the encoding format for the data product. This is done using the Undefined value of the S100\_DataFormat attribute in the discovery metadata of every compliant data product. See S-100 Part 4a for further details.

#### A-6.2.7 Data content and structure

This section describes the different type of data model elements that are used within a Product Specification. S-100 defines a variety of different elements, but it is up to the individual Specification to specify which elements are allowed to be used. The following is an example of how a subset of elements would be defined within a Product Specification.

*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.>

An additional option to describe the model elements included in a Product Specification 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 links it with the usage within the Specification. Figure A-6-4 below shows an example of this method, describing all information types in a specific Product Specification.

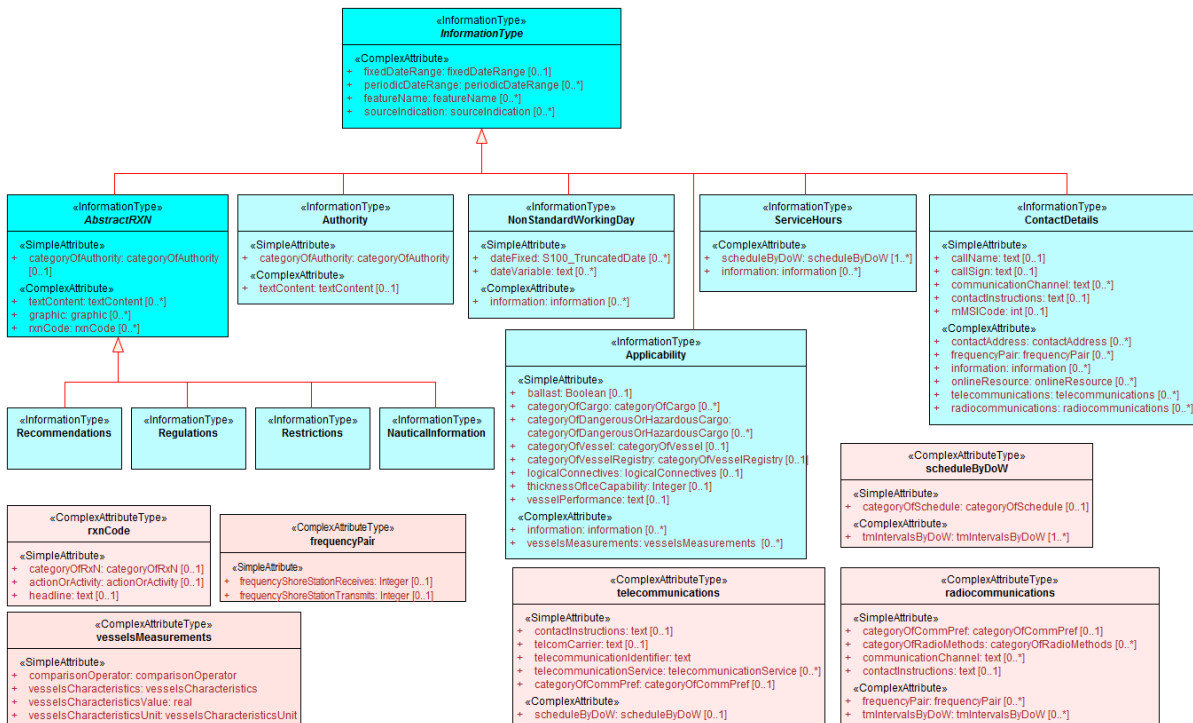


Figure A-6-4 – Example of an overview of S-Information Types, from S-123.

### A-6.2.8 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.

When a data model is too big to be fully reproduced in UML within in a Product Specification, the Feature Catalogue in combination with specific subsets of the overall UML Application Schema can serve as a substitute to the requirement of full Application Schema being described in the Product Specification.

Product Specifications that are endorsed by the IHO under the S-1XX numbering system must use the IHO’s Feature Catalogue Builder to ensure compliance to S-100. It is possible to create feature catalogues external to the IHO Feature Catalogue Builder (FCB); however, the Feature Catalogue should then be validated against the S-100 Feature Catalogue schema.

### A-6.2.9 Dataset

This section specifies rules for the dataset. Types of rules may be related to dataset size limitations or types of data coverage and if the data may or may not overlap each other.

### A-6.2.10 Dataset loading and unloading

If datasets have multiple representations at different scales, then a loading and unloading strategy should be considered. There should be sufficient details to give system implementers enough information to know how to create or load datasets in the correct manner. For example, any rules regarding dataset loading can be described using visual instructions like in Figure A- below.

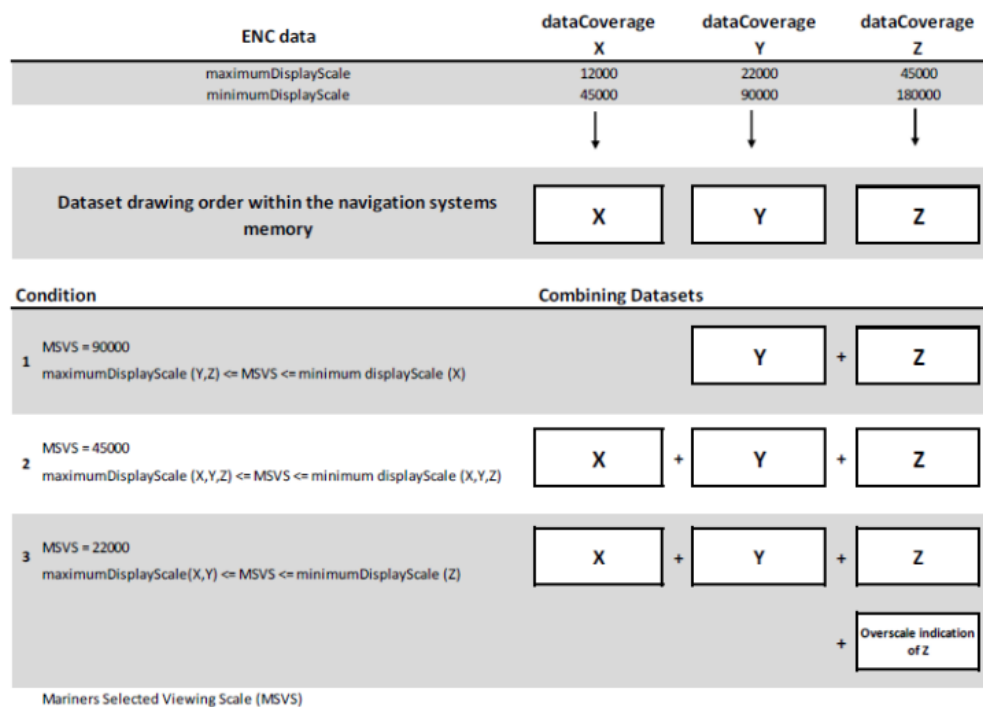


Figure A-6-5 – Example from S-101 data loading rules

### A-6.2.11 Geometry

This section describes the geometry rules that apply within a given Product Specification. For vector data the description should include which S-100 Level of Geometry is used, as well as any exceptions to the rules as stated in S-100 Part 7, clause 7-4.3. For coverage data, there should be a description of which spatial model and grid from S-100 Part 8 is used. Any grid, point set or TIN (Triangulated Irregular Network) structure rules and characteristics should be included, such as shown in the example in Figure A-6-6 below.

| Attribute                     | Value  | Remarks  |
|-------------------------------|--|--|
| <b>S100_Grid Coverage</b>     |  |  |
| dimension                     | 2  |  |
| origin                        | gridOriginLongitude,<br>gridOriginLatitude         | Values from Carrier Metadata   |
| axisNames                     | 'Longitude','Latitude'                             |  |
| offsetVectors                 | gridSpacingLongitudinal,<br>gridSpacingLatitudinal | Values from Carrier Metadata   |
| extent: low                   | '0, 0'   |  |
| extent: high                  | 'numROW-1, numCOL-1'                               | Values from Carrier Metadata   |
| sequencingRule: type          | 'linear'   |  |
| sequencingRule: startSequence | '0, 0'   |  |
| commonPointRule               | 'average'  |  |
| interpolationType             | 'discrete'   | There is no spatial interpolation for surface currents   |
| rangeType                     | name:data type                                     | Pairs which describes an attribute type included in the range of the coverage: e.g., 'surfaceCurrentSpeed: real' |
| <b>S100_Point Coverage</b>    |  |  |
| domainExtent                  | EX_GeographicExtent (ISO 19115)                    | Envelope based on all longitudes and all latitudes.  |
| axisNames                     | 'Longitude','Latitude'                             |  |
| rangeType                     | name:data type                                     | Pairs which describes an attribute type included in the range of the coverage                                    |
| metadata                      | URI  | Link to metadata   |
| commonPointRule               | 'average'  |  |
| geometry                      | GM_Point   |  |
| value                         | Real number  | Corresponds to speed and direction values  |

Figure A-6-6 – Example of attributes and their values for S100\_Grid Coverage and S100\_Point Coverage.

### A-6.2.12 Reference systems

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. A Product Specification may include more than one vertical reference system, such as one for sounding data and one for height data.

Part 6 of S-100 provides information on 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 Product 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. Figure A-6-7 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 A-6-7 – Example of Reference Systems references

As noted in the example in Figure A-6-7, the EPSG Register 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 includes an enumerated list named S100\_VerticalAndSoundingDatum. The most commonly used vertical and sounding datums are included in this list. The enumerated list can be extended by requesting the IHO S-100 Working Group to include additional values.

### A-6.2.13 Object identifiers

It is recommended that rules for persistent global identifiers for feature and information objects are included within a Product Specification. Identifiers may be omitted where the physical realities dictate otherwise or it is known that a reference to the object will not be needed. For example, identifiers need not be defined for cartographic objects.

Identifiers of instances should utilize the Maritime Resource Name (MRN) concept and namespace. The MRN namespace is administered by the International Association of Lighthouse Authorities (IALA) through the website <http://mrnregistry.org>, which also contains references to the full set of rules that apply to the MRN concept. The topmost namespace urn:mrn remains fixed, with subsequent name spaces separated by colons, and available through the application process explained on the website. Any organization wishing to issue MRN conformant identifiers should apply for a name space from IALA, or from an organization that already has a namespace registered, such as IHO. S-100 Part 11, Appendix 11-E contains additional details about the MRN concept.

Guidance should be included on persevering persistent global identifiers on objects throughout their lifecycle, including when they are reused in other products. Maintaining persistent global identifiers between products can help with interoperability and assist users and systems in identifying identical features between data products.

### A-6.2.14 Data quality

All S-100-based Product Specifications should include comprehensive ways of capturing information about the quality of the data. Part C of this Guideline includes comprehensive guidance on how to address aspects of data quality.

### A-6.2.15 Data capture and encoding instructions

Any S-100-based Product Specification should provide information on how data conforming to the Product Specification 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 real world examples to the data model. For example, the DCEG can explain how different types of underwater rock are to be encoded using a specific data model,

including which feature class should be used; what attributes; and their expected values that correspond to the different types of underwater rock.

The data capture guide is mostly used by the data producers and serves as a collective instructional document of globally common rules on how to create data according to a specific Product Specification. The document will grow with experience as more special cases get resolved into a globally agreed-upon 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.

#### **A-6.2.16 Maintenance**

Generally, data created from an S-100-based Product Specification will not remain valid indefinitely. Therefore, it is necessary to specify how data created in accordance with a Product Specification will be maintained, including the updating of datasets and support files.

There are two main types of updating routines:

- As needed: Datasets are updated when there is a need to do so, and are to be considered current information until there is a further update. Electronic Navigational Charts and Nautical Publications are two types of data that are generally maintained in this manner.
- By schedule: Datasets are updated on a fixed schedule or interval; and users can always anticipate when new datasets become available. Surface current and water level information are two types of data that are generally maintained in this manner.

Once updating routines have been established for products, it is necessary to establish the means by which to achieve these updates. Again, there are two main options: incremental updates; and whole dataset replacement, both of which are elaborated upon in the next two sections.

##### **A-6.2.16.1 Incremental update**

This type of updating method is when a previously issued dataset is partially updated by inserting new or modifying information. This method is useful where there are bandwidth considerations and the changes are relatively minor within the scope of the whole dataset.

An example is the addition of two features to a dataset that contains thousands of other features; the incremental update would then be a much smaller dataset that contains only the revision instructions to the main product data, or base dataset. Once the revision instructions are applied, the updated dataset will include the additional two objects.

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. This will include all the changes applied previously via incremental update, 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 4, ISO 8211 and HDF5 encodings support this type of updating. GML encoding does not yet support this type of encoding.

##### **A-6.2.16.2 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 forecasted data of a certain natural phenomenon is replaced with updated forecast data and the update data invalidates the replaced data by virtue of being more recent. All encodings in S-100 support this method of updating.



### A-6.2.16.3 Support file maintenance

Updating support files in S-100-based Product Specifications is done through the metadata that goes with the support files. The issue date and management information is included in the discovery metadata file that is part of the exchange set. Below is an example of instructions for updating support files.

*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 references 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.

### A-6.2.17 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 important where data is delivered to systems that include a level of data automation where standardized delivery structures may be automated. It is also useful to specify when data products are 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 (reproduced in Figure A-6-8 below), and details can be found in S-100 Part 4a, Appendix 4a-D.

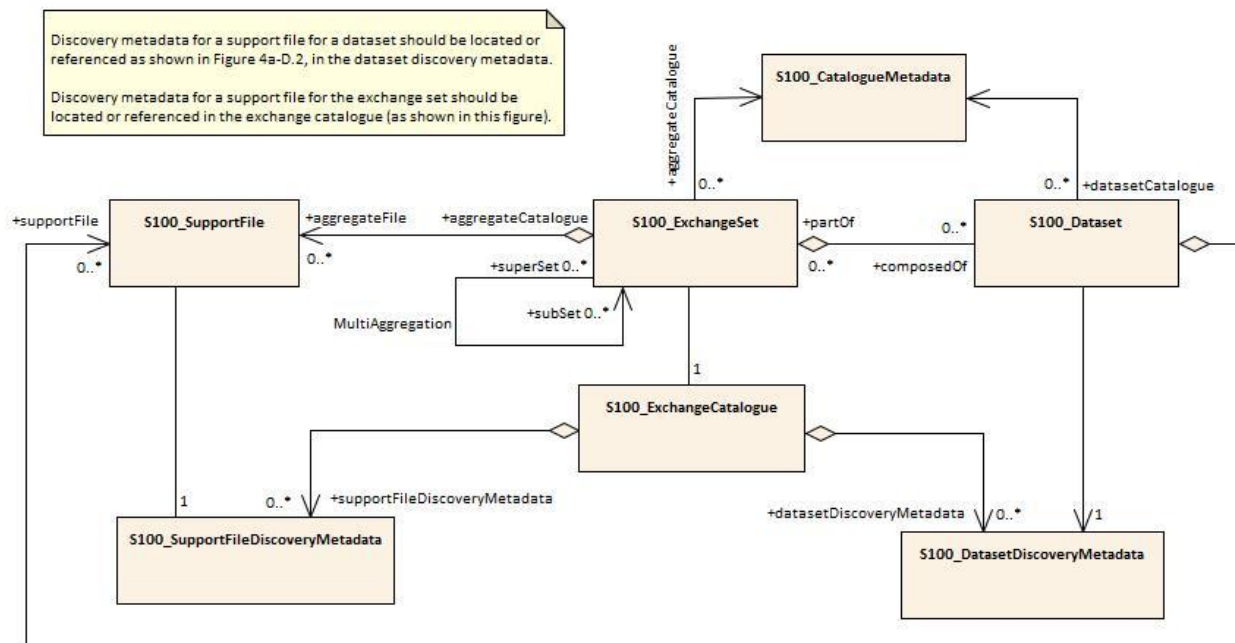


Figure A-6-8 – S-100 Exchange Set (S-100 Ed. 4.0.0 Figure 4a-D-3)

Within an exchange set, there are datasets; but it may also carry support files. There are two types of supporting files within the exchange set: support files for the individual dataset; and support files for the exchange set. Support files that are associated with the individual dataset usually include file types such as

text files and image files; while support files that are associated with the exchange set are usually 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.

#### **A-6.2.17.1 Services and data streams**

S-100 includes an alternative distribution method called online data exchange, which is described in Part 14. This method can be used for a set of data or data that have a continuous nature. The latter is also known as “streaming data” and is used where the circumstances requires a more dynamic information flow to be available, such as monitoring of ship movement in a VTS setting.

Online data exchange between applications or devices can follow different communication patterns to support the variety of maritime operational needs. Multiple clients can interact with a service to interchange data that is modelled with S-100 and can be distinguished between unidirectional message streams like AIS and interactive information exchange like a web feature service (WFS). Context for a communication can be given by using the concept of session-oriented communication, which is when the communication between distinguished communication partners is assigned to a logical entity – a session. This permits metadata to be defined for the interactions assigned to the session.

The means of communication for the use of a service should be defined in a communication stack. Specifying a communication stack ensures that communication for the service is harmonized and makes implementation easier.

The communication is organized by a stack as defined by the ISO-OSI Reference Model:

- Session protocols (for example, WSDL, SOAP, REST, SOS) to define message types;
- Encoding and compression (for example, GML, XML, ISO 8211, HDF, ...) to serialize data;
- Communication protocol (for example, HTTP) with encryption (for example, HTTPS) to define interaction between gateways; and
- Transportation Layer (for example, TCP/IP) with encryption (for example, SSL) to define the transportation node between gateways.

The stack is depicted in Figure A-6-9 below.

## ISO/OSI Layer

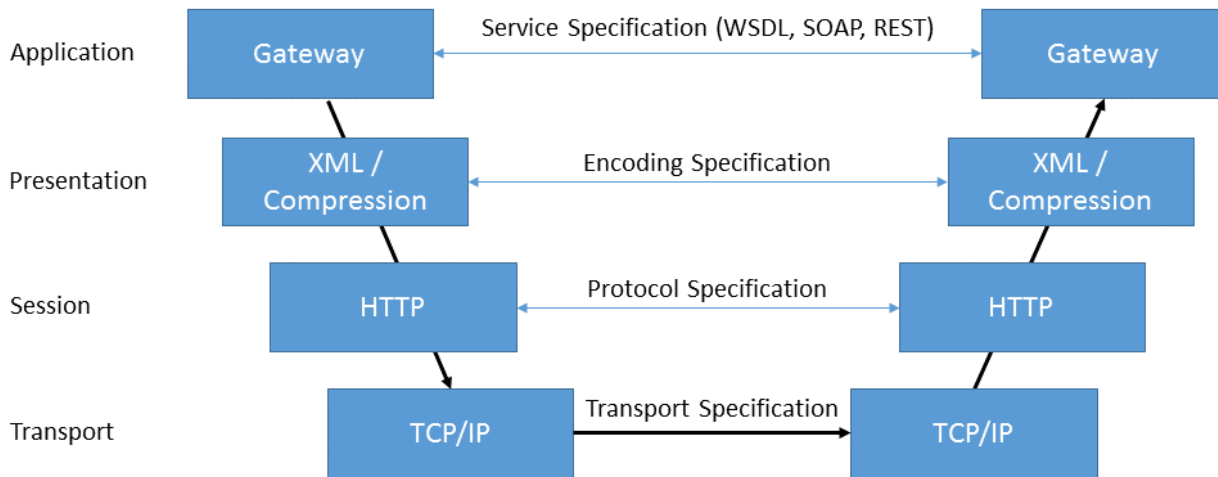


Figure A-6-9 – Communication Stack

S-100 Part 14 only addresses the concepts in the application and presentation layers. The lower layers are out of scope of S-100.

Product Specifications that use the online exchange method must describe the concepts that are used to structure a session; and must explain how the data is transferred within sessions. A session-oriented service typically contains three components, each handling different types of data:

- Session component: Describing the handling of the session data (service request, service response, login, login response, logout).
- Service component: Describing the information to maintain the service (for example keep alive messages, service status).
- Data component: Describing the data itself (for example, Vessel Traffic Image data (objects)).

Any Metadata required for each component should be detailed in the Product Specification.

#### A-6.2.18 Dataset naming rules

Dataset naming should follow a standard pattern to give implementers greater predictability of incoming datasets. All dataset naming conventions are recommended to follow these rules as much as possible.

XXXYYYYØØØØØØØØ

XXX is the product code (for example, 123 is for Maritime Radio Service; 101 for ENC)

YYYY is the producer code according to the Producer Code Register

ØØØØ is an arbitrary length unique code in alphanumeric characters

If useful, the Product Specification can include a differentiating character or code (for example, the underscore ( \_ ) character) in the 'ØØØØ' space of the file name.

Support files should follow a similar naming convention.

### A-6.2.19 Metadata

Metadata is data about data. In S-100 the primary purpose of metadata is to provide information about the identification; spatial and temporal extent; quality; application schema; spatial reference system; and distribution of digital geographic data. Metadata 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; etc. Figure A-6-10 below depicts some typical purposes and audiences for metadata.

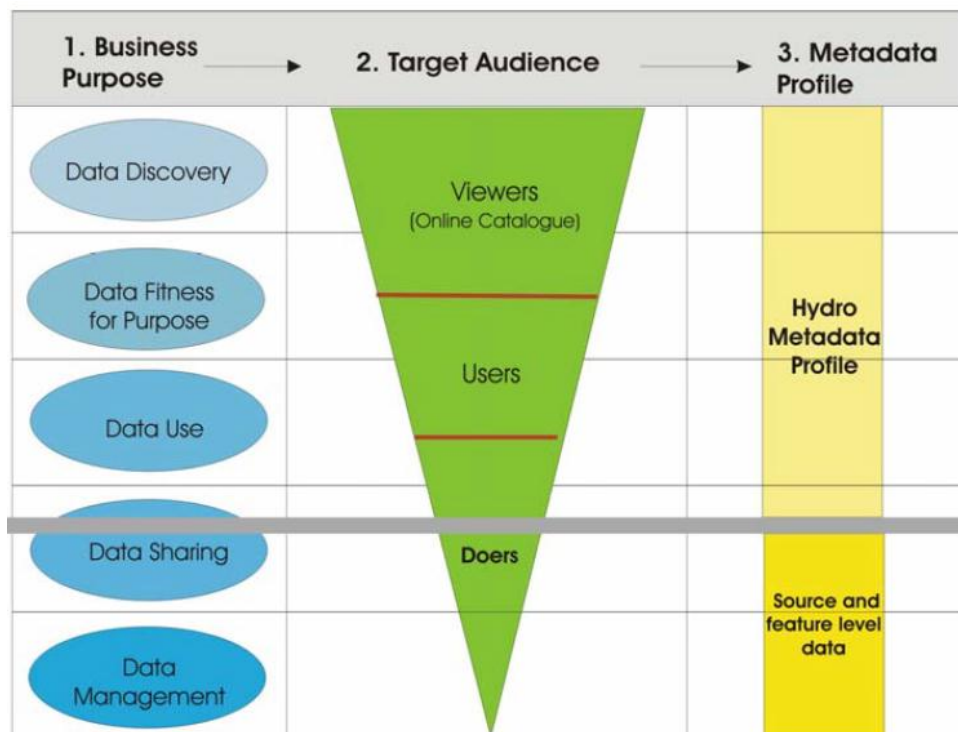


Figure A-6-10 – Example levels of metadata

S-100 Part 4 (Metadata) specifies the minimum metadata elements that must be included with a conformant data product. Moreover, S-100 requires that Discovery and Quality metadata are 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 Specification; and these should be defined using ISO 19115-1 and ISO 19115-3, with extensions or restrictions if required. S-100 Part 4, Appendix 4a-E contains the rules that apply when extending or restricting the minimum metadata.

The Product Specification Application Schema shall specify how metadata is packaged in conformant 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 (for example, 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 and it is highly recommended to stay within the S-100 metadata scope.

### **A-6.2.19.1 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. These are called Discovery Metadata in S-100 and they are used within the exchange set to enable users to learn about the content without having to open each dataset or support file.

#### **A-6.2.19.1.1 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 minimum metadata and extends it in a few places to provide more details (for example, about reference datums and issue dates of the data). More information about Discovery Metadata for datasets can be found in S-100 normative Appendix 4a-D, Discovery Metadata for Information Exchange Catalogues.

#### **A-6.2.19.1.2 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 minimum metadata and extends it to provide information about the management of support files in order for them to be updated. More details can be found in the S-100 normative Appendix 4a-D, Discovery Metadata for Information Exchange Catalogues.

#### **A-6.2.19.1.3 Metadata for streamed data and services**

Metadata for streamed data or services is described in the S-100 normative Appendix 4a-D, Discovery Metadata for Information Exchange Catalogues. S-100 Part 14 specifies additional and other metadata. As of Edition 4.0.0, Appendix 4a-D and Part 14 have yet to be reconciled. Product Specification developers should, at this time, use the metadata from Part 14; and according to need, supplement with metadata from Part 4, including using the principles of metadata extension detailed in Part 4, Appendix 4a-E.

### **A-6.2.20 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 done by including a Portrayal Catalogue. The Portrayal Catalogue is a collection of defined portrayal instructions for a Feature Catalogue; and includes portrayal functions, symbols and portrayal context.

Two types of Portrayal Catalogues are possible in S-100: XSLT and Lua. 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 (defined in SVG format); line styles, colours, etc; and makes 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 clause A-7.4 of this Guideline for more details).

Portrayal Catalogues can be provided, for example, 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 consideration should be given to efficient distribution and the aim 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.

Many of the IHO Product Specifications will be used in systems that have some degree of type approval requirements (for example, ECDIS). Instructions for the classification society conducting the type approval should be added to Product Specifications where appropriate. These instructions should include guidance on tolerances for minor deviations and definitions of what constitutes a minor deviation.

#### A-6.2.21 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. Such a section can be used to give context of intended use, or used to elaborate on special circumstances that impact stakeholders, and so forth.

## A-7 IHO S-100 infrastructure

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

### A-7.1 IHO Geospatial Infrastructure (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 online Registry engine called S-100 Geospatial Information (GI) Registry. The IHO GI Registry can be accessed at <http://registry.iho.int>. This Registry provides the facility to access and maintain the various S-100 Registers. The S-100 GI Registry contains subordinate Registers, depicted in Figure A-7-1 below.

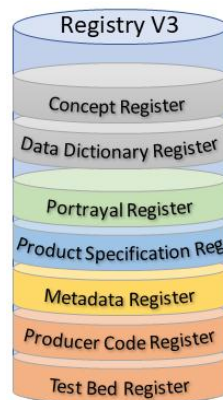


Figure A-7-1 – IHO GI Registry (Version 3)

Each Register type can be further sub-divided into Domains (for example, Hydrographic, Inland ENC, AtoN, etc). The administration of the IHO GI 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 versions);
- retired (no longer recommended for use); or
- not valid (proposed but not accepted or no longer acceptable).

Due to this classification and time stamps, a version of a Feature Catalogue references 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, a new version of a Product Specification is not automatically required as a consequence. The category of “not valid” items is included in the Registers specifically to help identify the inappropriate reintroduction of previously rejected proposals.

### A-7.2 Feature Catalogue Builder

A Feature Catalogue is a machine-readable expression of the data model for a Product Specification. It 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 in the creation of a Feature Catalogue for an S-100-based Product Specification. The software interacts with the IHO GI 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 A-7-2 below illustrates the concept of the IHO FCB.

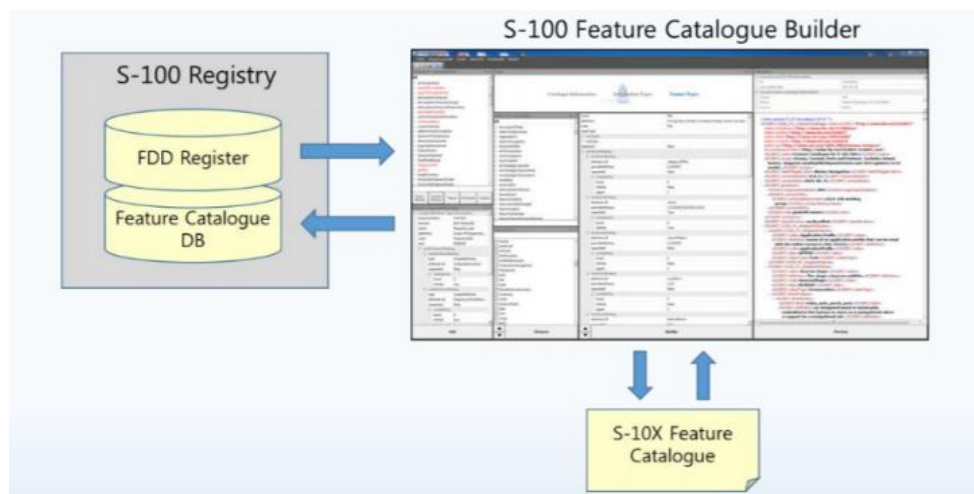


Figure A-7-2 - S-100 Feature Catalogue Builder

The FCB also contains a function for working with proposals that have yet to be added to the Registry.

### A-7.3 DCEG Builder

To simplify the creation of the Data Classification and Encoding Guide (DCEG) for a Product Specification, 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 A-7-3 below shows a high-level overview of the process to create a DCEG via the DCEG Builder. An Application Schema is required to make the Feature Catalogue, which equates to an XML representation of the Application Schema. The use of the Feature Catalogue can be described in the DCEG.

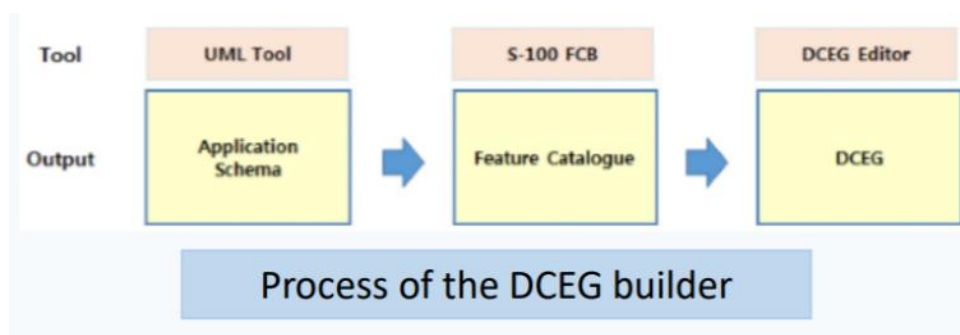


Figure A-7-3 – Process of creating a DCEG using the DCEG Builder

Use of the DCEG Builder is optional in the preparation of the Product Specification. Any S-100-based Product Specification must include sufficient instructions for how to encode information in a manner that conforms to the Specification, but these instructions do not have to follow a specific DCEG style. However, the DCEG style is simple to understand and by utilizing the DCEG builder, it is also easy to create tables<sup>2</sup> of feature and information types, attributes, associations and encoding instructions; and to ensure that these tables are consistent with the Feature Catalogue. The IHO-style feature tables (or equivalent) can be prepared and maintained using ordinary office word processing software, but experience shows that ensuring initial and continued conformance to the XML Feature Catalogue may be a significant task requiring much effort to maintain and keep current. Figure A-7-4 below shows how the DCEG Builder is connected to the IHO GI Registry; 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. It should be noted that the DCEG Builder is dependent on the Feature Catalogue being registered within the Feature Catalogue Database.

<sup>2</sup> A suggested format for such tables is described in the IHO S-100 Production Specification Template, which is referenced in S-100 Part 11, Appendix 11-D and available from the IHO web site.



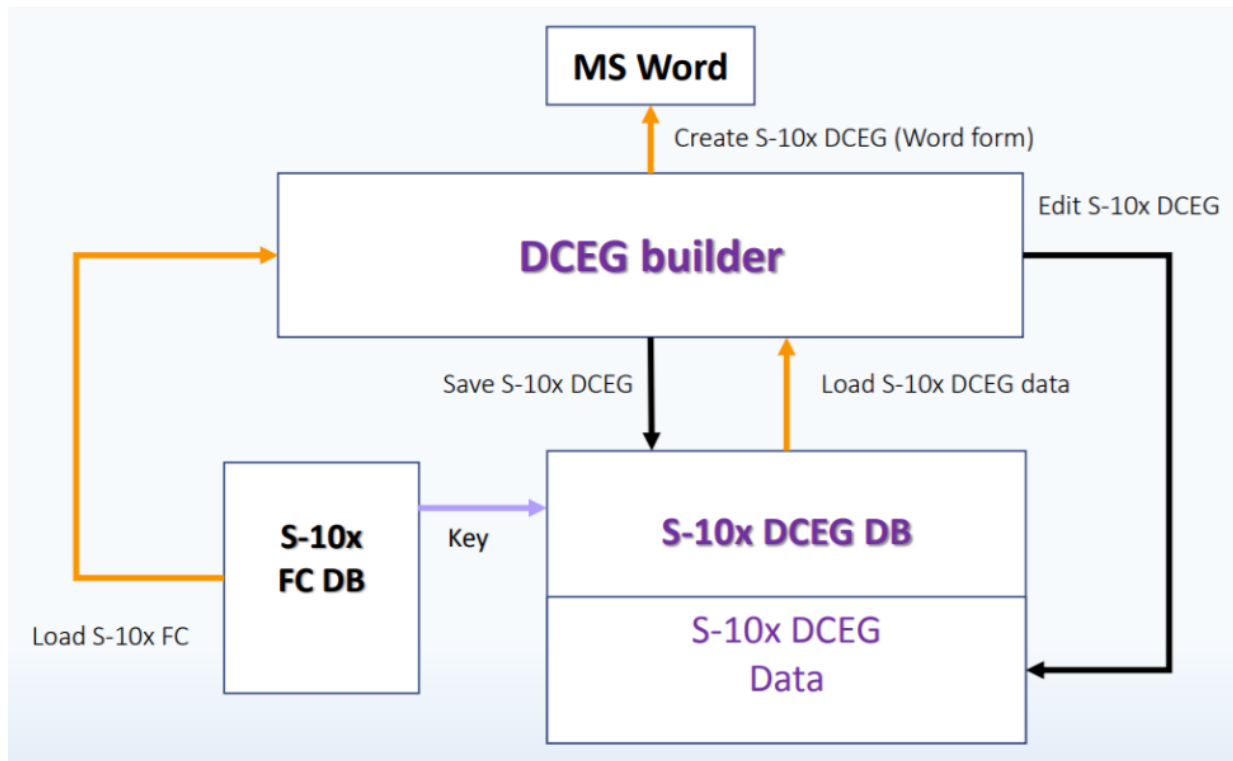


Figure A-7-4 – DCEG Builder architecture

### A-7.4 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). In either case, they must comply with the structure specified in S-100 Part 9 and the S-100 Portrayal Catalogue Schema. The IHO infrastructure includes a PCB for XSLT for any Product Specification development team wishing to utilize it in the creation of a Portrayal Catalogue for an S-100-based Product Specification.

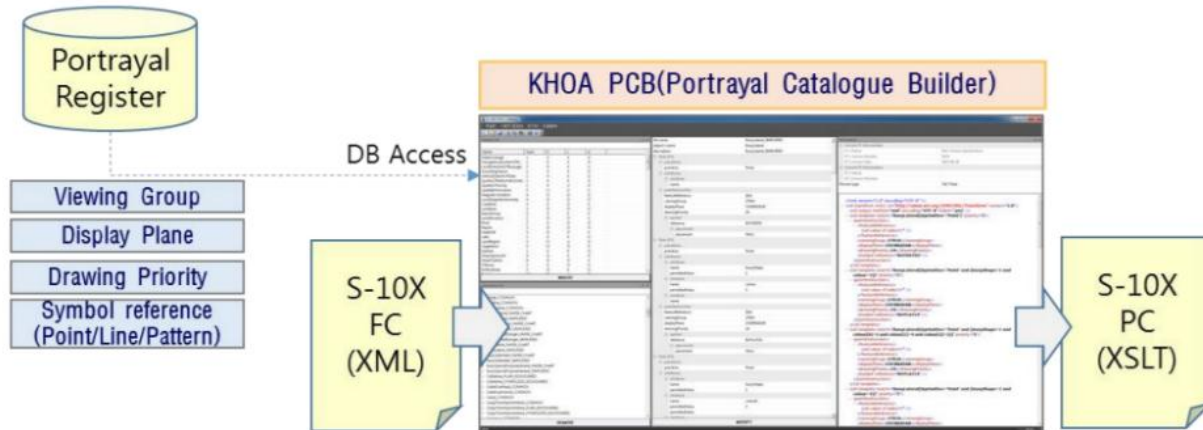


Figure A-7-5 - Portrayal Catalogue Builder

The software interacts with the Portrayal Register in the IHO GI Registry and the Feature Catalogue and provides an interface for binding elements available in the IHO GI Registry together to form symbols, line styles and area patterns for the desired elements in the Feature Catalogue. Figure A-7-5 above illustrates the PCB concept.

A PCB for Lua Portrayal Catalogues is planned for the future.

## Part B - Steps in the Development of a Product Specification

### B-1 Overview

S-100 the Universal Hydrographic Data Model is a hydrographic geospatial data standard that can support a wide variety of hydrographic-related digital data sources, and is fully aligned with mainstream international geospatial standards, in particular the ISO 19000 series of geographic standards. This alignment enables easier integration of hydrographic data and applications into geospatial solutions. S-100 is inherently more flexible than S-57 and makes provision for such things as the use of imagery and gridded data types, enhanced metadata and multiple encoding formats. It also provides a more flexible and dynamic maintenance regime for features, attributes and portrayal via a dedicated on-line registry. S-100 provides a framework of components that enable the building of standardized Product Specifications for the modelling of hydrographic data, thus providing true interoperability between different data standards and systems.

S-97 is a Guideline intended for developers and maintainers of product specifications that utilize the IHO framework standard S-100 (Universal Hydrographic Data Model).

### B-2 Introduction

The intent of Part B is to explain the various components of a Product Specification for those that are developing S-100 based Product Specifications. It describes the typical phases of development for the typical Product Specification; and a prescribed order for each phase. It provides guidance on how the rules contained in S-100 are used to build the components of a Product Specification, including documentation components; UML models; and machine-readable components such as XML Feature Catalogues and data formats.

### B-3 References

- |             |   |
|-------------|---|
| ISO 8211    | Specification for a data descriptive file for information interchange structure implementations. ISO/IEC 8211, 1994.  |
| ISO 19109   | Geographic information – Rules for application schema. ISO 19109:2005. (Since replaced by ISO 19109:2015, but S-100 Edition 4.0 still references the 2005 edition.) |
| ISO 19115-1 | Geographic information – Metadata – Part 1 – Fundamentals. ISO 19115-1, 2014, as amended by Amendment 1, 2018.  |
| ISO 19115-2 | Geographic information – Metadata - Part 2 – Extensions for imagery and gridded data. ISO 19115-2, 2009.  |
| ISO 19115-3 | Geographic information – Metadata - XML schema implementation for fundamental concepts. ISO/TS 19115-3, 2016.   |
| ISO 19119   | Geographic Information – Services. ISO 19119, 2016.   |

|                                    |  |
|------------------------------------|--|
| ISO 19136                          | Geographic information – Geography Markup Language (GML). ISO 19136, 2007. (Also available as OGC 07-036 Geography Markup Language (GML) Encoding Standard. Open Geospatial Consortium Inc., 2007.)                  |
| ISO 19157                          | Geographic information – Data Quality. ISO 19157, 2013, as amended by Amendment 1, 2018.   |
| ISO 19757-3                        | Information technology – Document Schema Definition Languages (DSDL) – Part 3: Rule-based validation – Schematron.   |
| IHO Resolution 2/2007 (as amended) | Principles and Procedures for Making Changes to IHO Technical Standards and Specifications.  |
| S-52                               | Specifications for Chart Content and Display Aspects of ECDIS / Specifications pour le Contenu Cartographique et les Modalites D'affichage des ECDIS.  |
| S-58                               | IHO S-58 – ENC Validation Checks, Edition 6.1.0, September 2018.   |
| S-99                               | IHO S-99 – Operational Procedures for the Organization and Management of the S-100 Geospatial Information Registry, Edition 1.1.0, November 2012.  |
| S-99A                              | IHO S-99, Annex A – Conventions and Guidelines for the Content of the IHO GI Registry. (Under development).  |
| S-100                              | IHO S-100 – Universal Hydrographic Data Model Edition 4.0.0, December 2018.  |
| S-101                              | IHO S-101 – Electronic Navigational Chart Product Specification, Edition 1.0.0 (draft), July 2018.   |
| S-111                              | IHO S-111 – Surface Current Product Specification, Edition 1.0.0 (draft), June 2018.   |
| S-122                              | IHO S-122 – Marine Protected Areas, Edition 1.0.0, January 2019.   |
| S-123                              | ISO S-123 – Marine Radio Services, Edition 1.0.0, January 2019.  |
| SVGTiny                            | Scalable Vector Graphics (SVG) Tiny 1.2 Specification. W3C Recommendation 22 December 2008. URL: <a href="http://www.w3.org/TR/2008/REC-SVGTiny12-20081222/">http://www.w3.org/TR/2008/REC-SVGTiny12-20081222/</a> . |

NOTE: In this document, “S-100” means S-100 Edition 4.0.0 unless a different edition is explicitly identified.

## B-4 Terms and abbreviations

### B-4.1 Terms

#### abstract class

an object class which cannot be **instantiated**, or is designated in an information model as not allowed to be instantiated

NOTE: Subclasses of an abstract class may be either abstract or non-abstract.

#### aggregation

special form of association that specifies a whole-part relationship between the aggregate (whole) and a component part (see **composition**) [ISO 19103]

#### application

manipulation and processing of data in support of user requirements [ISO 19101-1:2014]

**application schema**

**conceptual schema** for data required by one or more **applications** [ISO 19101-1:2014]

**association**

semantic relationship between two or more classifiers that specifies connections among their instances [ISO 19103]

**attribute**

(1) named property of an entity [ISO/IEC 2382-17:1999]

NOTE: Describes a geometrical, topological, thematic, or other characteristic of an entity.

(2) UML: feature within a classifier that describes a range of values that instances of the classifier may hold [ISO/IEC 19501:2005 (Adapted)]

(3) XML: name-value pair contained in an element [ISO 19136]

**base64**

an encoding designed to represent arbitrary sequences of octets in a form that allows the use of both upper- and lowercase letters but that need not be human readable [IETF RFC 4648 (restyled)]

**code**

representation of a label according to a specified scheme [ISO 19118:2011]

**codelist**

value domain including a code for a permissible value [ISO 19136]

**codespace**

rule or authority for a code, name, term or category [ISO 19136]

EXAMPLE: Examples of codespaces include dictionaries, authorities, codelists, etc.

**composition**

form of aggregation association with strong ownership and coincident lifetime as part of the whole [ISO 19103]

**conceptual model**

**model** that defines concepts of a universe of discourse [19101-1:2014]

**conceptual schema**

formal description of a **conceptual model** [ISO 19101-1:2014]

**data client**

an end-user receiving encrypted S-100-based data; the data client will be using a software application (for example ECDIS) to perform many of the operations detailed within the S-100 protection scheme [S-100 Part 15 (adapted)]

EXAMPLE: An ECDIS user.

**data permit**

file containing encrypted product keys required to decrypt the licensed products, normally created specifically for a particular **data client** [S-100 Part 15 (adapted)]

**data server**

an organization producing encrypted data files or issuing **data permits** to **data clients** [S-100 Part 15, (adapted)]

**feature**

abstraction of real world phenomena [ISO 19101:2003]

NOTE 1: A feature may occur as a type, class, or an instance. Feature type or feature instance should be used when only one is meant. **Feature class** should be used in the context of a model or Application Schema.

EXAMPLE The phenomenon named 'Eiffel Tower' may be classified with other phenomena into a feature type 'tower'.

**feature association**

**relationship** that links instances of one feature type with instances of the same or a different feature type [ISO 19110]

**feature attribute**

characteristic of a feature [ISO 19101]

NOTE A feature attribute type has a name, a data type and a domain associated to it. A feature attribute instance has an attribute value taken from the value domain of the feature attribute type.

EXAMPLE 1: A feature attribute named 'colour' may have an attribute value "green" which belongs to the data type "text."

EXAMPLE 2: A feature attribute named 'length' may have an attribute value "82.4" which belongs to the data type "real."

**feature catalogue**

a catalogue containing definitions and descriptions of the **feature types**, **feature attributes** and **feature associations** occurring in one or more sets of geographic data [ISO 19110]

**feature class**

a class in an **application schema** or **model** that represents a **feature**.

**feature type**

an element in a **feature catalogue** that describes a **feature**, its **attributes**, and **associations**.

**identifier**

a linguistically independent sequence of characters capable of uniquely and permanently identifying that with which it is associated [adapted from ISO/IEC 11179-3:2003]

**information type**

an identifiable unit of information in a dataset with only thematic attribute properties [S-100 3-5.1.2 (adapted)]

**instantiate**

represent by a concrete instance [Merriam-Webster Online <https://www.merriam-webster.com/dictionary/instantiate>]

**interface**

named set of **operations** that characterize the behaviour of an entity [ISO 19119:2005]

**metadata**

information about a **resource** [ISO 19115-1]; data that defines and describes other data [ISO 11179-3:2013]

**model**

abstraction of some aspects of reality [ISO 19109-2015]

**operation**

specification of a transformation or query that an object may be called to execute [ISO 19119:2005]

NOTE: An operation has a name and a list of parameters.

**register**

set of files containing identifiers assigned to items with descriptions of the associated items [ISO 19135]

NOTE: Descriptions may consist of many types of information, including names, definitions and codes.

**registry**

information system on which a **register** is maintained [ISO 19135]

**relationship**

semantic connection among **model** elements [ISO 19103]

**resource**

identifiable asset or means that fulfils a requirement [ISO 19115-1]

EXAMPLES: Dataset, dataset series, service, document, initiative, software, person or organization.

**scheme administrator**

organization solely responsible for maintaining and coordinating the protection scheme specified by S-100 [S-100 Part 15 (adapted)]

**service**

distinct part of the functionality that is provided by an entity through **interfaces** [ISO 19119:2005]

**spatial object**

object used for representing a spatial characteristic of a feature [ISO 19107:2003]

**stream**

in online data exchange: a continuous sequence of fragmented data to be transported by a communication system [S-100]

**universe of discourse**

view of the real or hypothetical world that includes everything of interest [19101-1:2014]

**vocabulary**

terminological dictionary which contains designations and definitions from one or more specific subject fields [ISO 1087-1:2000]

## B-4.2 Abbreviations

|       |  |
|-------|--|
| CORBA | Common Object Request Broker Architecture                                  |
| COTS  | Commercial Off-The-Shelf   |
| CRS   | Coordinate Reference System  |
| CSS   | Cascading StyleSheets  |
| DCEG  | Data Classification and Encoding Guide                                     |
| ECDIS | Electronic Chart Display and Information System                            |
| ECS   | Electronic Chart System  |
| ENC   | Electronic Navigational Chart  |
| FC    | Feature Catalogue  |
| GFM   | General Feature Model  |
| GI    | Geospatial Information   |
| GML   | Geography Markup Language  |
| HSSC  | IHO Hydrographic Services and Standards Committee                          |
| IALA  | International Association of Lighthouse Authorities                        |
| IEC   | International Electrotechnical Commission                                  |
| IETF  | Internet Engineering Task Force  |
| IHO   | International Hydrographic Organization                                    |
| IMO   | International Maritime Organization  |
| ISO   | International Organization for Standardization                             |
| JCOMM | WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology |
| MRN   | Maritime Resource Name   |
| OEM   | Original Equipment Manufacturer  |
| RENC  | Regional ENC Coordinating Centre   |
| SVG   | Scalable Vector Graphics   |
| W3C   | World Wide Web Consortium  |
| XML   | eXtensible Markup Language   |
| XSD   | XML Schema Definition  |
| XSL   | eXtensible Stylesheet Language   |
| XSLT  | XSL Transforms   |

## B-5 Overview of steps for development

When developing an S-100 based Product Specification it is recommended that developers consider each of the stages defined in the basic development process. The stages are outlined below and described in greater detail in the following sections.

### B-5.1 Basic development process

The steps described below apply to the development of a Product Specification for a new data product.



- 1) **Initiation.** Identify the need for a new data product; define its scope; and decide the boundaries between the new product and existing data Product Specifications. Obtain sample source material. Describe typical application use cases.
- 2) **Develop the Domain Model/Application Schema.** Define the classes and attributes that describe the domain and which are relevant to the data product. Define the relationships between the classes and specify applicable constraints. Prepare one or more UML diagrams describing the Domain Model.
- 3) **Populate the IHO GI Registry from the Domain Model.** Propose amendments to existing classes and attributes and propose new classes and attributes for addition to the Concept and Data Dictionary Registers in the IHO GI Registry using the Registry interface. Follow up on any returned proposals or queries from the Registry Manager or Domain Control Body.
- 4) **Develop the Feature Catalogue.** Prepare the XML Feature Catalogue from the feature and information classes, attributes and relationships as approved in the IHO GI Registry, utilizing the Feature Catalogue Builder.
- 5) **Define data transfer modes and packaging.** Determine whether data products are to be delivered as data files contained in transfer (exchange) sets, by web services (and if so, identify or outline a service protocol), e-mail, etc. Determine whether data is to be delivered in real or near real-time. Identify constraints and requirements arising from delivery mechanisms and communication constraints such as message size, bandwidth limitations, availability of communications to customers, licensing and payments, encryption, etc.
- 6) **Define metadata.** Survey the metadata elements listed in S-100 for their appropriateness to the data product and its allowed packaging and delivery methods. Define appropriate values and restrictions for the metadata elements listed in S-100. Consider whether additional product-specific metadata elements are needed.
- 7) **Define the data format.** Select an appropriate data format. S-100 provides for 3 standard delivery formats (ISO 8211, GML, and HDF5). Define format-imposed items (for example embedded header metadata). Prepare format-specific artefacts if necessary (for example GML “application schema” XSD files for the GML format).
- 8) **Develop the Data Classification and Encoding Guide (DCEG).** The DCEG is intended for cartographers, editors, and data encoders, rather than application developers or OEMs; and should be written from that perspective. A DCEG should contain enough overview and general material about basic concepts such as data types, features, information types, associations, etc, to give its intended audience a basic grounding in the concepts they will need to apply; but its main focus should be on what to encode in the data product, and how to encode it.
- 9) **Define portrayal symbols and rules.** Determine the symbols to be used for portrayal and the rules for generating displays from the data product. This means deciding which features should be displayed; which combinations of attribute values or associated data objects should be assigned distinct symbols; and which pre-existing symbols and colours can be reused and which need to be defined. It also includes defining the new elements in the appropriate formats.
- 10) **Registration of portrayal elements.** Propose any new portrayal components (for example symbols, colour tokens, line styles, area fills, etc) to the Portrayal Register in the IHO GI Registry using the Registry interface. Follow up on any returned proposals or queries from the Registry Manager or Domain Control Body.
- 11) **Develop the Portrayal Catalogue(s).** Prepare a Portrayal Catalogue (or Catalogues) for the features and information types which are intended to be displayed in the intended application domain(s) and usage scenario(s). This consists of encoding the rules in the appropriate portrayal language using the symbols determined in step (8).

- 12) **Define the spatial reference system.** Identify the recommended coordinate reference system and vertical datum(s).
- 13) **Define data product packaging and maintenance.** Define the content and structure of delivery packages, updating of data, and any auxiliary content delivered either with or as an adjunct to data.
- 14) **Define validation checks and quality measures.** Define tests for the spatial, structural, and conceptual integrity of datasets. Define format-specific implementations of validation checks (for example Schematron rules [ISO 19757-3] for the GML format). Define quality measures for the datasets (quality measures will generally be statistics about passed and failed tests).
- 15) **Determine interoperability** with other data products. Determine which if any product groups in Interoperability Catalogues are supplemented or enhanced by the data product; and how the IHO Interoperability Catalogue will be affected by the new product, including updates to display priorities, interleaving, predefined combinations and other interoperability rules and operations.
- 16) **Prepare sample data** for test-beds. Create sample datasets and exchange sets conforming to the data format, packaging, and Feature Catalogue defined in the Product Specification.
- 17) **Testing and feedback.** Carry out tests of data production and use of the sample data in selected applications to validate the correctness, completeness, consistency, and utility of the Product Specification, including related artefacts such as the Feature Catalogue and XML schemas.

Development of Product Specifications is an iterative process of continuous refinement as it moves through the S-100 Readiness Levels. Developers should expect to cycle through the stages above multiple times, in smaller or larger cycles depending on experiences and results of each stage. For example, the first preparation of sample data (Stage 16) might identify gaps in the Domain Model and require revisiting Stage 2 (development of the Domain Model); or the development of portrayal rules (Stages 9 - 11) may identify gaps in the Domain Model (Stage 2). Passage through some intermediate stages in the cycle may be trivial on the second and later passes, as the results may be unaffected by discoveries during later stages in the process – for example, packaging and delivery or choice of spatial reference systems are likely to be unaffected by changes identified during the preparation of sample data.

Figure B-5-1 below depicts the development process, including probable cycles.

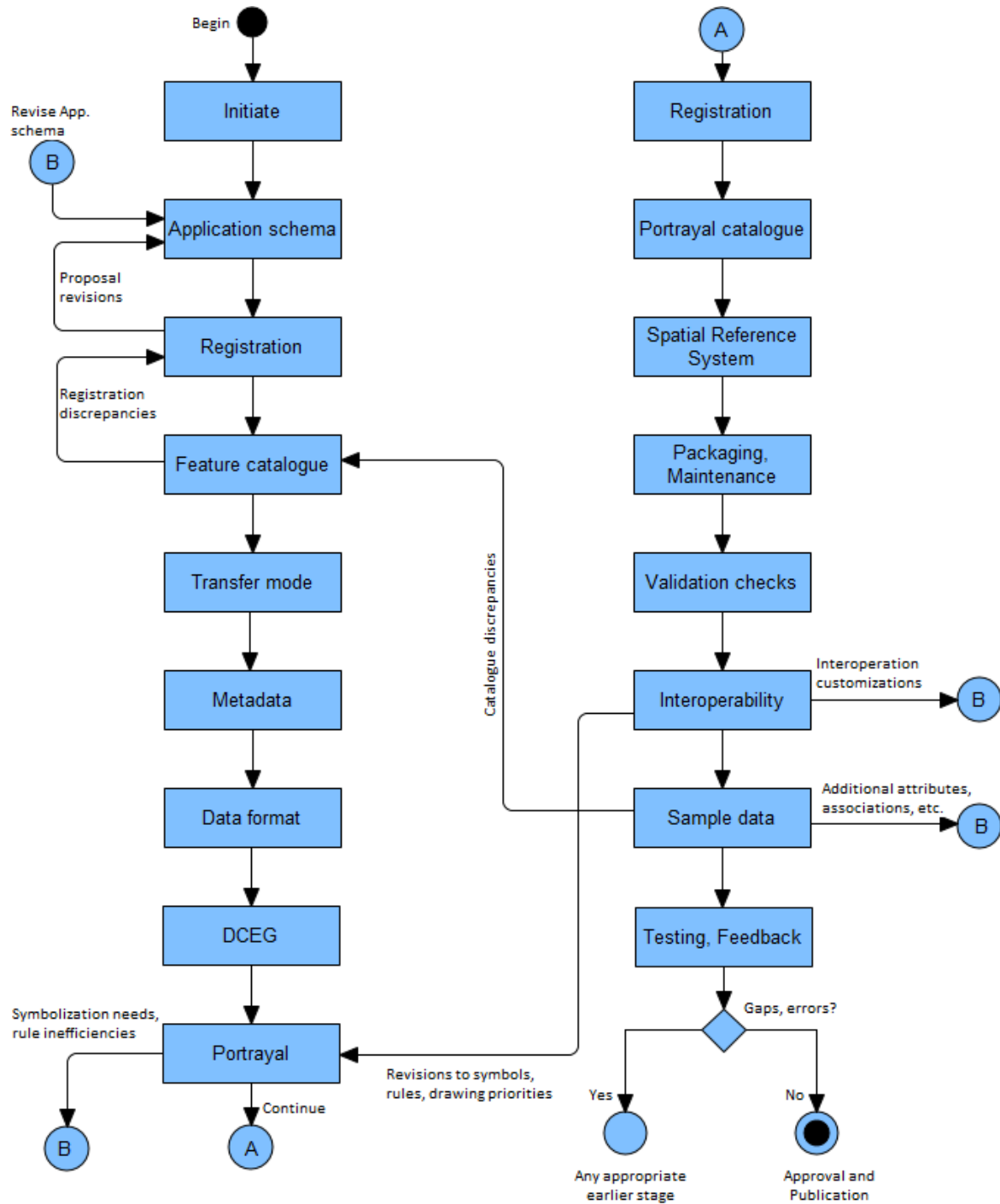


Figure B-5-1 – Product Specification development process

S-100 Part 11, Appendix 11-A describes an idealized process (Figure 11-A-1 – Product Specification process) for the basic mechanical steps in developing a domain model plus related registry actions. The figure also includes black-box stages for “Coordinate Reference System” and “Product Specification Documentation” at the end of the process. The actual process is an iterative refinement process.

## B-5.2 Review cycle for S-100-based Product Specifications

IHO Resolution 2/2007 (as amended) outlines the process for development and approval of IHO based Product Specifications. As part of this lifecycle it is important that Working Groups and Project Teams engage with the IHO Data Quality Working Group (DQWG) to review Product Specifications for adherence to S-100 data quality aspects and completeness. While this review can occur prior to Edition 1.0.0 (Testing and Implementation Edition), the DQWG will review and provide feedback during the Edition 1.0.0 to 1.9.n phase Specification development. Edition 1.0.0 of any IHO S-100-based Product Specification is not designed for regular use in approved arrangements or for provision of operation services by purpose. Edition 2.0.0 of a Product Specification is considered to be mature and can be used for implementation. Figure B-5-2 below shows the IHO Review Cycle as included in IHO Resolution 2/2007.

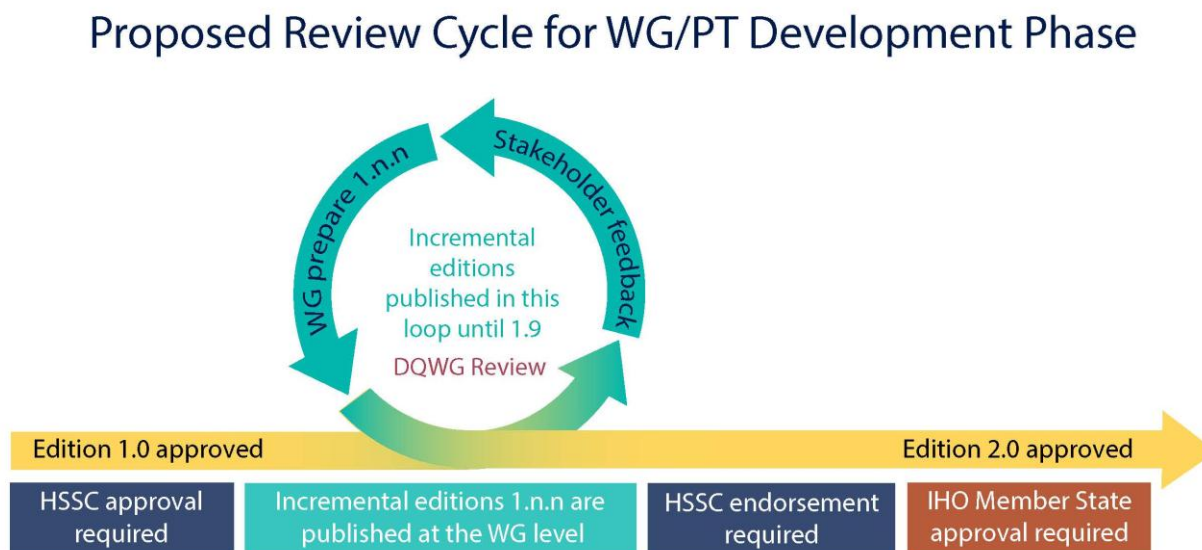


Figure B-5-2 – Review Cycle (IHO Resolution 2/2007 (as amended))

## B-5.3 Process for extending a Specification

If there is a demonstrated need to extend an existing Product Specification, developers should review the steps in the basic process in section 5.1 and determine which steps can be omitted.

For example, if features, information types, attributes, or relationships are added, removed, or modified, most of the steps in the basic process will be needed. The spatial reference system and packaging/maintenance may not change in such a revision and the corresponding phases can be skipped.

If the extension consists of defining a new transfer mode, such as adding a streaming delivery mode for data products originally intended for exchange set delivery, it is likely that only the following stages need be executed:

- Transfer mode – to add the streaming mode for transfer, define specifications for servers, etc.
- Metadata – to define which metadata elements in the exchange set mode apply to the streaming mode, whether any new metadata elements are needed, and how metadata is delivered or made available in the new mode.

- Data format – to consider whether the already defined data format is appropriate for streamed transfer and compatible with the intended service delivery and protocols; and specify an appropriate format if necessary.
- Packaging and maintenance – to specify the appropriate data wrapper formats and service protocols; for example WFS, REST, etc. Defining new wrapper formats or protocols is permitted but should be the last resort.
- Sample data – preparation of sample data for testing the streaming mode.
- Testing and feedback – this should include a prototype or simulated streamed data transfer environment.

## **B-6 Initiation**

The initiation phase consists of the following activities. Except for the first (identifying the need) most of the other activities can be done in parallel.

- 1) Identify a need that a data product will fill – this will normally be the result of external circumstances such as definition of a service portfolio; application-driven demand for new kinds of information; new legal requirements for shipping; IMO decisions; etc.
- 2) Define the scope of the product – the subject area, the kind of information it is expected to contain; and equally important, what information it will not contain.
- 3) Determine if existing Product Specifications can be extended. If so, such an extension will probably consume less time and effort than developing a new Product Specification.
- 4) Determine sub-areas within the product; that is, the scopes within the Product Specification or a new scope for an existing Product Specification.
- 5) Define constraints – domain, application, platform, etc.
- 6) Collect samples of source information. These generally include existing databases; and official, unofficial, government and commercial publications – especially those in wide use.
- 7) Define application use cases.
- 8) Outline application functionality enabled by the data product.
- 9) Define delivery modes (transfer set, messages, web services, etc).
- 10) Obtain approval from the appropriate sponsoring organization.
- 11) Put together a Project Team.

If the Product Specification development is performed under the umbrella of the IHO then the following step also needs to be included in the initiation stage:

Before deciding to include a new S-100-based Product Specification in the HSSC Work Program, the following considerations should be taken into account:

- 1) Has the need for the Product Specification proposed been documented?
- 2) Has a compelling need been demonstrated with detailed examples?
- 3) Has the Product Specification been considered within the scope of the HSSC?
- 4) Has the description of the proposed Product Specification sufficiently addressed the cost to the IHO and possible legislative and administrative burdens?

## 5) Has the development duration been considered?

To simplify the above-mentioned decision making process, the Product Specification description should provide a set of information beforehand to HSSC for consideration, as shown in Table B-6-1 below.

| Product Specification description |  |   |
|-----------------------------------|--|---|
| Action                            | Description  | Comments  |
| Product Specification number      | The intended number for this Product Specification   | Keep this section empty until HSSC have assigned a number   |
| Title                             | The title of the product specification   |   |
| Abstract                          | A brief summary of the Product Specification summarizing:<br>1. the intended use<br>2. the primary and secondary user<br>3. the expected functionality |   |
| Product Specification Scope       | The overall scope of the Specification   | This should include the expected SOURCE(s) of the data.<br><br>For example: NIPWG publications...a source would be the Coast Pilot/Sailing Directions for S-126 (that way if someone thinks Tide Tables should be here instead of with S-104 it will be listed) |
| Justification                     | The reason why this Product Specification should be developed  |   |
| Specification Interoperability    | Any interoperability with other Product Specifications within the S-100 family   | Specify the limits of the Product Specification   |
| S-98 Applicability                | Applicable to S-98 (Yes or No)   | S-98 Interoperability Specification in S-100 Navigation Systems   |
| Cooperation with other HSSC WGs   | Specify which WG will be involved to which extend  | NCWG: Portrayal<br>DQWG. Specific information on Data Quality   |
| Budget                            | Statement of budget need and the figure  | Is IHO budget needed? Will the development be financed by an external party and to what extent?<br><br>There should be general guidelines for how to calculate this...with reasonable values based on reality and evidence                                      |

| Product Specification description |  |   |
|-----------------------------------|--|---|
| Schedule                          | Description of the intended time frame | Or at least the steps to accomplish since this is very difficult to determine |

*Table B-6-1 – Product Specification description template*

## **B-7 Develop data model (Application Schema)**

### **B-7.1 Introduction**

The Application Schema as defined in S-100 is usually synonymous with “domain model”. It is a specification of the classes, attributes, and relationships relevant to the data product.

A Product Specification will include only a single Application Schema (which may be broken up into multiple diagrams). However, in theory a Product Specification that describes different scopes may need to distinguish Application Schema for different scopes – for example a product that includes both vector and coverage data might need two Application Schemas.

To minimize complexity, Product Specification developers should try to avoid defining multiple Application Schemas.

Note that this does not prevent a single Application Schema from being depicted using multiple diagrams.

The Application Schema should describe only features, information types, their attributes and relationships that are part of the data product. Any other classes, constraints, or elements that are used to further understand the data product and its role in applications, services or service portfolios should be distinguished from the Application Schema and documented separately in the Product Specification.

Model developers should try to limit the number of model elements while still allowing implementations to make appropriate (conceptually appropriate, logically appropriate, consistent, correct and performance-based) distinctions.

The principles of data normalization learned in relational database design should be kept in mind; but model developers should also note that an S-100-based domain model and Application Schema are feature-attribute-relationship models and not database designs.

### **B-7.2 Steps in model development**

- 1) Determine whether the data product is coverage or vector data. Coverage data is characterized by values of characteristics distributed over an area or areas, while vector data is characterized by localized regions (points, and/or areas) that possess boundaries and do not exhibit internal variation in characteristic values (or where such internal variations can be ignored).
- 2) Identify the concepts in the application domain. This will involve reviewing the source material to identify important features and information in the domain which will be useful to end-users. Source material will include the sample texts identified in the initiation phase and, if available, documentation and data dictionaries of relevant applications, requirements for existing applications, related standards and circulars from IEC, IMO, IALA, etc.

- 3) Search for existing concepts using key words (classes, attributes and relationships) in the IHO GI Registry which can be re-used.
- 4) Develop new concepts only for those that do not yet exist in the IHO GI Registry. This will involve examining the source material mentioned earlier in more detail to pin down concepts and their definitions. The process for submitting proposals is described in S-99.
- 5) Define the classes and attributes that describe the domain and are relevant to the data product. If classes, attributes and codelists/enumerations are already defined in the IHO GI Registry or existing Product Specifications they can be reused. If not, the Project Team will need to develop and define classes and attributes, including listed values for enumeration and codelist types.
- 6) Define the relationships between the classes. Relationships should be defined in order to capture those that exist in the real world and to make links which are useful for application processing. Both reasons will often apply.

EXAMPLE 1: A structure/equipment relationship between classes modelling structure objects and classes modelling equipment mounted on the structure.

EXAMPLE 2: A contact information relationship between classes modelling pilot service areas and contact information for pilot services available in that area.

- 7) Specify any constraints applicable to the classes, attributes, and relationships. Examples are constraints requiring conditional encoding of attributes, exclusive relationships (that is, when an instance is allowed to participate in only one of multiple possible relationships), etc. Generally, structural restrictions are depicted in UML class diagrams, while value restrictions on individual attributes are not (to reduce clutter). Whether depicted or not, any restrictions should be enforced and documented in the appropriate section or artefact of the Product Specification.
- 8) Prepare one or more UML class diagrams describing the domain model. Recommended practices for S-100 models are based on ISO TC211 recommended practices as modified by clause B-7.6 of this document.
- 9) Prepare supporting text explaining the overall structure of the Application Schema; and for each diagram explaining the purpose of each diagram and the relationships between the classes. In addition, explain any additional subtleties of the classes or their relationships that may not be obvious or should be specially noted. This text should not be a listing of classes, attributes, and relationships; but instead should clarify the purpose of the model depicted in the diagram by explaining what domain phenomenon the diagram captures and how the classes and relationships express it. For example, from S-122:

*Some protected areas require reports to be filed with authorities when certain events occur such as an animal strike or pollution event. Other areas require reporting to specified authorities when entering, leaving, etc. These requirements are modelled by association of a ShipReport class to the Authority class. The area in question is modelled by a feature of the requisite type; for example., a MarineProtectedArea or VesselTrafficServiceArea, as described in clause 6.2.1.3 [of S-122]. Any time requirements or constraints on the filing of the report are described by the noticeTime attribute, with explanations, if any provided in text form in the textContent attribute of ShipReport. Required reporting formats, if necessary to be included, are also described in the textContent attribute. [Figure ] shows the model elements that are used to carry*



these conditions, note that not all permitted associations or roles are included, in order to reduce clutter.

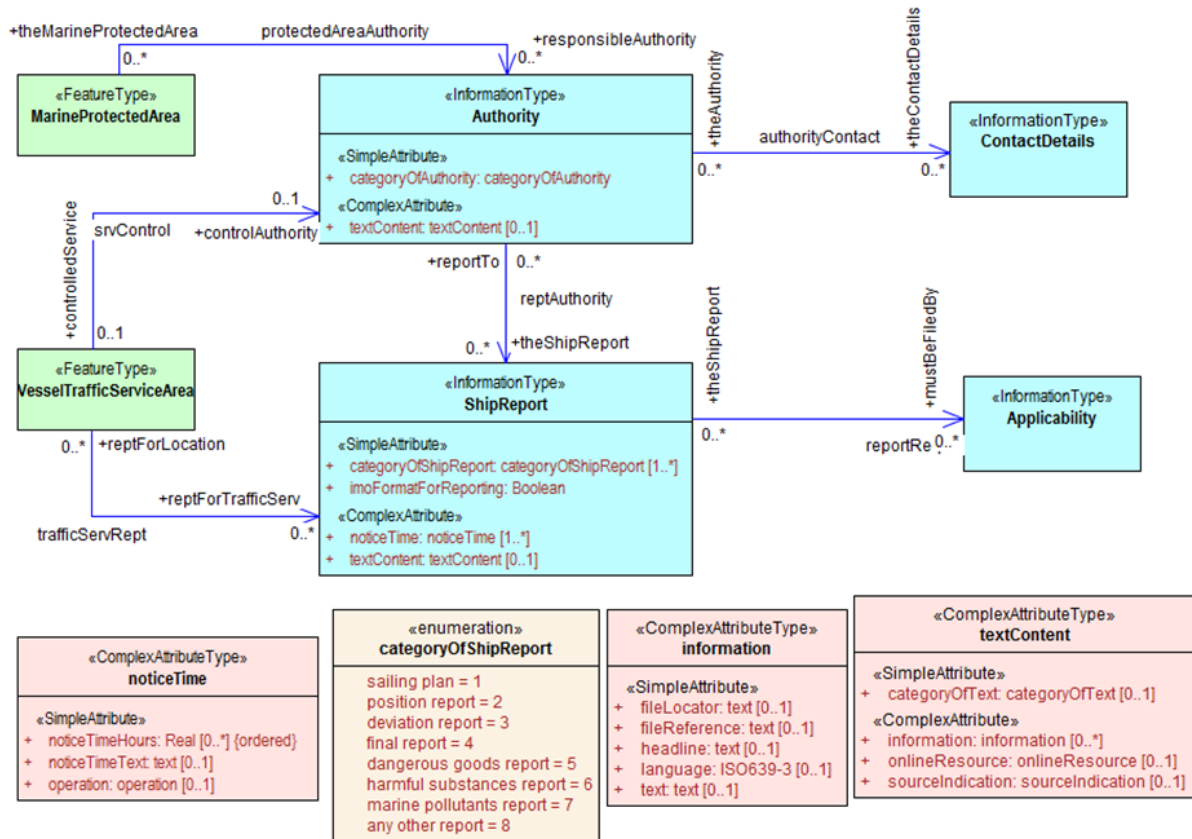


Figure B-7-1 – UML diagram depicting part of the S-122 Application Schema (from S-122)

### B-7.3 Relationship to the General Feature Model

The General Feature Model is a conceptual model for features, their characteristics and associations; and acts as the basis for the structure of Feature Catalogues. Feature and information types in the Application Schema must be realizations of the meta-classes S100\_GF\_FeatureType and S100\_GF\_InformationType from the S-100 General Feature Model (GFM) (S-100 Part 3, Figure 3-1); or subclasses of a class that realizes the appropriate meta-class. Attributes must be realizations of the thematic or spatial attribute meta-classes defined in S-100 Part 3, Figure 3-2; or subclasses of a realization. Product Specifications may define local root classes from which all their feature and information classes are derived, as shown in Figure B-7-2 below; or may realize feature and information type classes from S100\_GF\_FeatureType and S100\_InformationType.

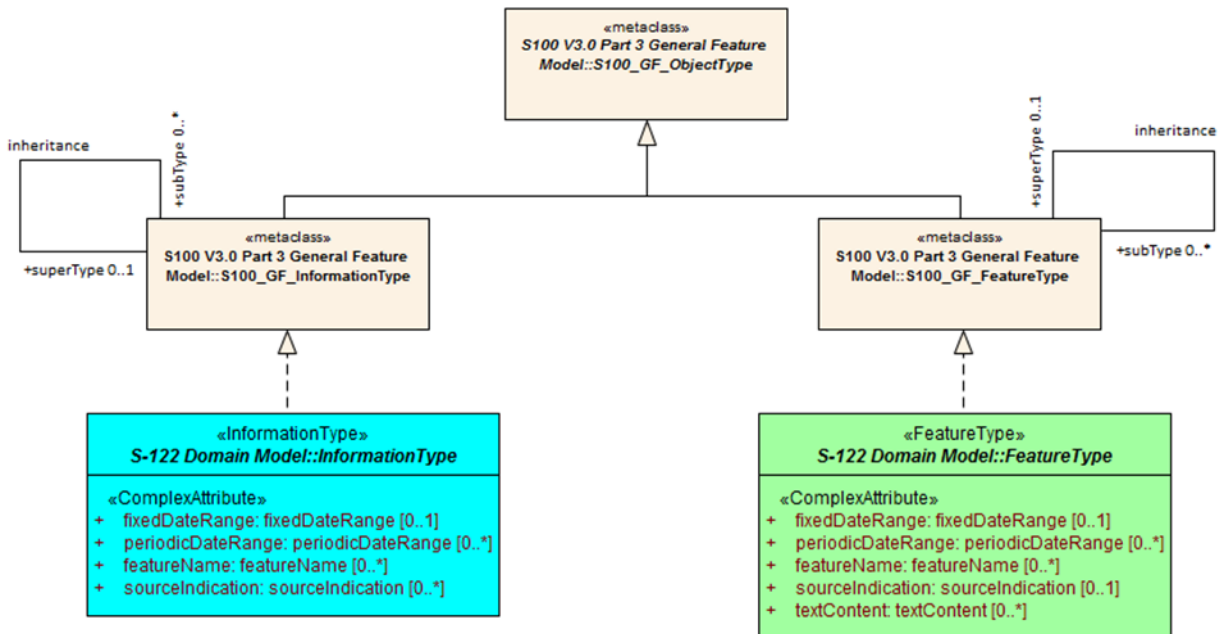


Figure B-7-2 – Example of realization from S-100 GFM

## B-7.4 Rules for Application Schemas

S-100 rules for Application Schemas are based on ISO 19109. The S-100 rules for Application Schemas are defined in S-100 Part 3, clauses 3-6 and 3-7.

### B-7.4.1 Application Schemas for vector data

Application Schemas for vector data should follow the guidelines as described below. S-100 Part 3, clause 3-6 describes the rules for Applications Schemas in greater technical detail:

- Features, information types, and complex attributes must be modelled as classes.
- Relationships are modelled by UML associations or association classes (the latter only when the association itself is characterized by attributes – see clause B-7.5.4.2 in this document).
- Attributes are modelled by UML attributes in the appropriate class.
- Associations must be labelled (have association names). Navigable association ends must also be labelled (should have role names).

NOTE: Diagrams may suppress depiction of labels for clarity and to reduce clutter. Labels may be defined by specific rules given in the Product Specification text instead of the UML diagram (for example a Product Specification is allowed to ‘label’ an association end using a statement like “The role of *FeatureX* in all its associations is *theFeatureX*” (see S-100 Part 3, clause 3-5.4.5 on default names for association ends).

- Spatial attributes must be modelled either as attributes with data type one or more (that is, union) of the allowed spatial types in the spatial schema, or an association between the class that represents a feature and one of the spatial objects defined in the spatial schema.
- Enumeration types and their listed values must be modelled by UML enumerations; codelists must be modelled as UML classes with tags specified in S-100 Part 3, clause 3-6.7.

- Standard schemas (for example the spatial schema, Feature Catalogue schema) must not be extended within Application Schemas.
- All classes used within an Application Schema for data transfer shall be instantiable. This implies that the integrated class must not be stereotyped <<interface>>.
- A UML Application Schema must be described within a UML package, which must carry the name of the Application Schema and the version stated in the documentation of the package.

#### B-7.4.2 Application Schemas for coverage data

The rules for Application Schemas for coverage data are described in S-100 Part 3, clause 3-7; and Part 8.

The rules are similar to the rules for vector data, with the following difference:

- Spatial types for coverage features are modelled by the appropriate point set, grid, or TIN type defined in S-100 Part 8.

#### B-7.4.3 Additional rules

Names of features and information types must use their camel case codes.

Vector feature classes must use the stereotype <<FeatureType>> and information classes must use the stereotype <<InformationType>>. A stereotype allows designers to extend the UML model by creating new model elements.

Coverage *type* elements (describing the coverage geometry) compliant to S-100 should use the appropriate stereotype from S-100 Part 8; and Application Schemas for coverage data may depict the data attributes by defining a <<FeatureType>> element with the thematic data attributes and associating it with the coverage type element.



Figure B-7-3 – Example of coverage feature and type elements conforming to S-100 (from S-111)

If necessary, Product Specifications may use domain-specific stereotypes in addition to the standard stereotypes.

Abstract classes are indicated by italicizing the class name (Enterprise Architect does this automatically if the “Abstract” checkbox is checked in the UI).

S-100 states that “the use of multiple inheritance shall be minimized, because it tends to increase model complexity”. Multiple inheritance is the situation where a class has more than one immediate superclass. Application Schema developers should note that multiple inheritance contravenes the S-100 GFM, which allows feature and information types to have at most one super-type.

## B-7.5 Other conventions and recommendations

### B-7.5.1 Reuse and harmonization

Before new elements are defined the IHO GI Registry should be checked for existing elements that can be re-used. The following should be also considered:

- Features, information types, and attributes should be re-used whenever possible.
- Structure and associations should be harmonized with S-101 and other existing related or complementary products.
- Defining similar but slightly different items should be avoided unless absolutely essential.
- Extensions such as additional listed values in an enumeration can be proposed to the IHO GI Registry, but conflicts such as different definitions for the same terms must be avoided if at all possible.
- Existing items may be reused with the addition of product specific constraints, such as limiting the set of allowed values for an enumeration or codelist type. Such harmonization includes, for example, re-using complex attributes defined in other Product Specifications with restrictions that exclude some of their sub-attributes.

### B-7.5.2 Features and information types

A feature class is “an abstract representation of real-world phenomena” (S-100 Part 3, clause 3-5.1.1). The Application Schema should define a feature class for describing:

- 1) A concept whose instances have a spatial (geographic) location. Such a class represents a **geographic** feature.
- 2) A concept whose instances are collections of the above. The collection may consist of instances of one such class or several such classes. Depending on the nature of the collection, the concept will be either an **aggregation** feature or a **composition** feature. In Application Schemas they are treated like geographic features but may or may not have a spatial attribute.
- 3) Cartographic information (such as feature names, labels, compass roses, legends) that are intended to be positioned at a specified location (or re-positioned from its default position relative to a feature) to avoid obscuring other features. Though not a representation of any real-world phenomenon, such information is also modelled as a feature and treated as a feature in the Feature Catalogue. Such features are called **cartographic** features. The information to be displayed may be encoded in an attribute of the cartographic feature class or identified by reference to another feature class.
- 4) Meta-information or metadata pertaining to all features (or defined subsets or defined attributes of features) in a particular area. These are called **meta** features.

The S-100 Feature Catalogue model provides a *featureUseType* element for designating feature types as “geographic”, “cartographic”, or “meta”. Aggregation and composition feature types are designated as being of “geographic” *featureUseType*.

Information types are identifiable objects that can be associated with features or other information types, in order to carry information particular to the associated objects. Information types may be considered classes that are typically used to share information among features and other information types. Information types have only thematic attribute properties.

### B-7.5.3 Superclasses and subclasses

Defining abstract superclasses is recommended when 3 or more conceptually similar classes exist in the model. The similar classes have some of the same attributes or relationships; and the allowed values of shared attributes are the same. It is not necessary that the classes bind exactly the same sets of attributes or have exactly the same relationships.

Subclasses inherit both associations and attributes, unless explicitly overridden. While overriding is permitted, it introduces additional complexity in data formats and implementations. Instead of overriding, the model should be changed to move the attribute or association down in the hierarchy and assign it only to classes where its use is permitted.

#### B-7.5.3.1 Superclasses versus category attribute

There are different ways of modelling the same item. For example, in order to indicate different buoy types, either the subclass approach or the enumeration approach (categoryOf) can be used. Both approaches are depicted in Figure B-7-4.

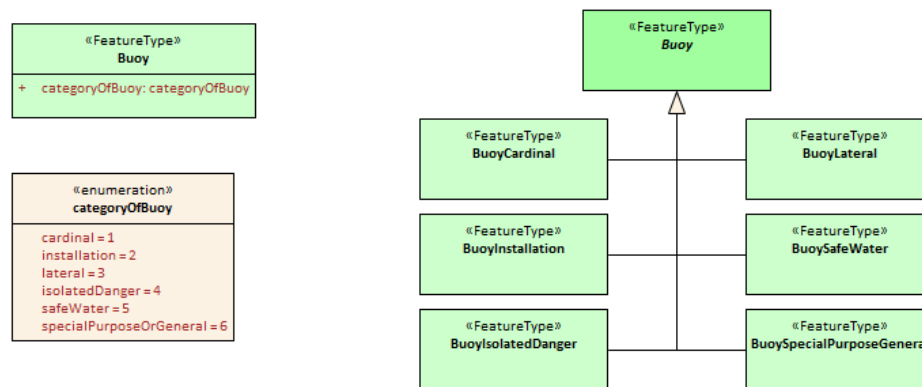


Figure B-6-2 – Illustration of alternative models using category attribute and sibling subclasses

It is up to the Product Specification development team to determine which approach best suits their needs. In order to determine this, consideration should be given to the following issues:

- 1) Will either approach result in a significant divergence from some external source?
- 2) Will the “category attribute” approach cause issues for portrayal because the symbols for the different types are presumably different? Or will the subclass approach result in unnecessary portrayal rules because the symbols are the same?
- 3) Will any of the subclasses have its own specific attributes or relationships? If so, the sub-class approach is preferable.
- 4) Are the different categories/subclasses in (or likely to be placed in) different viewing groups, or have different drawing order? If so, there is a slight preference for making subclasses. (Only “slight” because the portrayal rules and Interoperability Catalogue can use attribute values in assigning viewing groups to feature instances.
- 5) If subclasses are used, will that introduce situations where it may be necessary to encode coincident objects with different categories? The answer “Yes” suggests a preference for the categoryOf... approach.
- 6) Are the subclasses conceptually very different? “Yes” implies the sub-classes approach.

- 7) Which approach is likely to be compatible with external resources like existing databases and implementations?
- 8) If there are a large number of subtypes, then the categoryOf... approach may be preferable because it leads to more compact representations in UML diagrams and more compact DCEGs. (“Large” is obviously subjective, but will generally be between 5 and 9 based on research into human cognitive psychology and probable implementation methods in user interfaces – there will be variations dependent on concept semantics and similarities.)
- 9) Overall complexity of the Application Schema and Feature Catalogue. Sibling subclasses of features (or information types) generate more artefacts and documentation than a category attribute. They certainly mean an additional table for each subclass in the DCEG; an additional XML element for each in the Feature Catalogue; and a box in the UML diagram Application Schema for each class. To that extent sibling subclasses are a greater cognitive burden on encoders and developers.

#### **B-7.5.4 Associations and association classes**

##### **B-7.5.4.1 Navigability, source and target**

Association navigability should be indicated if the association is navigable in only one direction; that is, the model designer expects applications to access one object from the other, but not vice versa. Feature/feature and information/information associations are usually navigable in both directions, while feature/information associations must be navigable from the feature end but are not required to be modelled as navigable in the other direction. UML regards navigability information in UML diagrams as hints to implementations rather than hard requirements; and implementations and data formats are free to implement navigability in the most efficient manner.

Unidirectional navigability will normally also determine the source and target of the association.

An association’s source and target should be grammatically and semantically compatible with the name and definition of the association; for example, for the association Person/subscribes/Magazine the source should be Person and the target Magazine.

For feature/feature associations both ends should be named; for feature/information associations the information end should be named and the feature end may be named. This is an S-100 requirement, not a UML requirement.

##### **B-7.5.4.2 Association classes**

Association classes are a means of adding parameters (characteristics) to associations, rather than to the classes at the end of an association. An attribute of the association class characterizes the relationship between the classes at the ends of the association.

The use case for association classes is basically “whenever a relationship is characterized by one or more attributes”.

**EXAMPLE 1:** A specified set of vessels is COVERED by a regulation and another set of vessels is EXEMPT from the regulation. The sets of vessels are described by an information type class; the regulation by another information type class; and the relationship between them by an association class which has an attribute characterizing the relationship as inclusion or exclusion (of the specified subset in the specific regulation). This relationship is modelled by the **InclusionType** association class in Figure B-7-5 below.

**EXAMPLE 2:** Vessels with specified cargo and dimensions must use a specified pilot boarding place; vessels of smaller dimensions are recommended to use the boarding place; and warships are exempt from using

the pilot boarding place. The sets of vessels are described by an information type class; the pilot boarding place is a feature class; and the relationship between them by an association class which has an attribute stating whether the specified set of vessels is required/recommended/exempt from use of the pilot boarding place. This relationship is modelled by the **PermissionType** association class in Figure B-7-5 below.

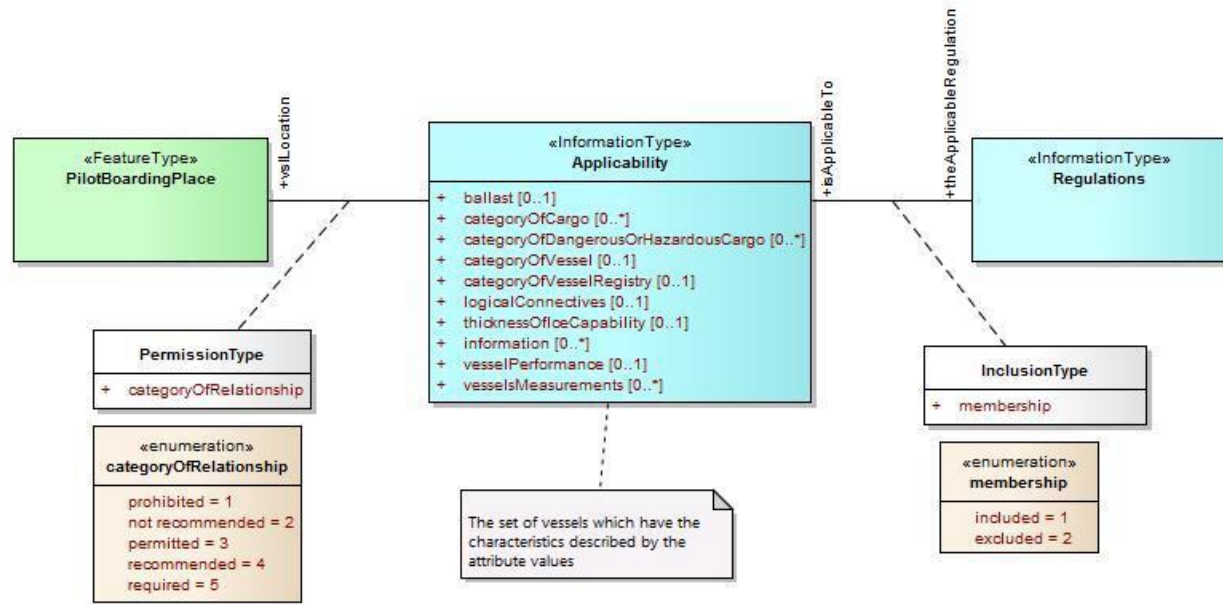


Figure B-6-5 – Examples of modelling with association classes

### B-7.5.5 Attributes in general

Certain attributes may use the same set of listed values as other attributes. For example, an enumeration for compass points may be shared by the attribute **windDirectionCompassPoint**, which is the direction for where wind is coming from; and **directionOfMovement**, which is the attribute describing where a weather system is going toward. At present this can be simply handled in the attribute's definition.

Complex and spatial attributes can be modelled as either named attributes in the UML model class element with a type corresponding to the spatial primitive or the name of the complex attribute; or alternatively, separate model elements linked to feature/information class by an association (ordinary association for spatial type, composition for complex attributes). The two methods are illustrated in Figure B-7-6 below. The second method is not suitable for complex models due to the additional boxes and association lines.

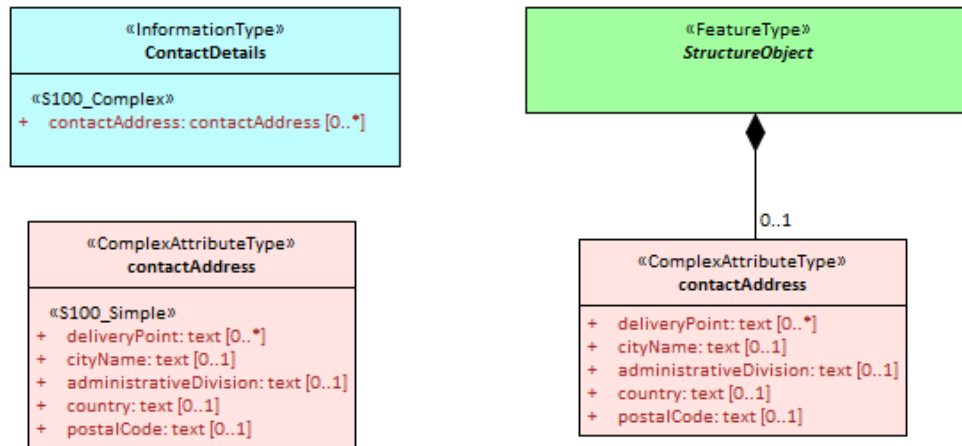


Figure B-6-6 – Methods of representing complex attribute bindings

### B-7.5.6 Codelist and enumeration attributes

Codelists should be used only when an enumeration is either unusable or inefficient (for example, if the full list of values is not known to the Product Specification authors or the list of allowed values is long, volatile, controlled by another authority and/or shared by multiple domains).

Individual members of a codelist or enumeration (“listed values”) are generally meaningful only in a given context (or “container”), which is a specific (named) enumeration. The context corresponds to an attribute concept.

### B-7.5.7 Labels and definitions for listed values

Labels should be short but informative, keeping in mind that both end-users and encoders may leverage them. Implementers of end-user display and production tools are likely to use listed value labels as ‘tooltips’ or explanatory text, or as the ‘display text’ of the attribute numeric code value for end-users. End-users are more likely to see labels than full definitions due to other demands on their attention and screen display constraints.

### B-7.5.8 Data types

S-100 defines a set of primitive and derived data types in a Table included in Part 2a, clause 2a-4-10. Attribute values of thematic attributes should be one of the types listed in this Table. If restrictions on the values are needed, Product Specifications may define constraints, if possible encoded using one or more of the elements in S100\_CD\_Constraints (length, range, pattern, and precision). Constraints which cannot be encoded must be documented as a ‘Remark’ or note.

The data types defined in S-100 can in principle be extended by Application Schemas but if this is done the Product Specification must define the extended data type in terms of the predefined data types in S-100 and use the predefined data type in the Feature Catalogue. Data formats may use their equivalent built-in types which are defined in the underlying format standard (for example HDF5 and XML built-in types) in order to leverage standard data validation software, provided the equivalence is documented either in the Product Specification or the underlying format standard.



### **B-7.5.9 Codes for listed values**

Numeric codes *must* be positive integers; and *should* be in the range 1-254 if possible (to allow data formats and implementations to use compact representations – however codes up to 65535 are allowed). Codes used for retired listed values can only be used if the proposal is a revision (supersession) of the retired listed value.

The IHO GI Registry also allows specification of alias and camel case codes for listed values. The camel case field should be completed with a camel case code derived from the label, beginning with a lowercase letter. The guidelines for camel case codes are specified in S-99 Annex A.

## **B-7.6 Recommended practices**

### **B-7.6.1 Reviews of model elements and structure**

Models should be reviewed at regular intervals while under development, with reviews involving domain experts as well as information modelling experts.

### **B-7.6.2 Diagram layout**

Common ‘best practices’ for layout of UML diagrams should be followed. In particular, diagrams should not contain too many elements; should minimize line crossings; and use vertical layouts for hierarchies (or left-right horizontal layouts if a vertical layout does not work). Lines representing associations should minimize the use of curved segments.

### **B-7.6.3 Colour coding of model elements**

Colour coding should be used to distinguish diagram elements for features; information types; enumerations and codelists; complex attributes; association classes; and constraints and notes. Abstract types should be indicated by darker shades.

Figure B-7-1 and Figure B-7-5 illustrate the use of colour coding to depict different kinds of UML elements. Compare the shades of the non-abstract feature and information classes in these figures to the abstract feature and information classes in Figure B-7-2.

S-100 departs from ISO TC211 recommendations for the use of black-and-white-only UML diagrams in order to distinguish between feature and information types (the concept of information type is unique to S-100).

### **B-7.6.4 Documentation tables**

These may be formatted like the UML schema documentation tables in S-100 or generated by the UML software. Whichever method is used, the documentation must document the classes, attributes, enumeration and codelist types; and associations in the Application Schema, including names, definitions, multiplicities, data types and roles.

### **B-7.6.5 Recommended software tools**

The S100 Working group recommend using Enterprise Architect™ to develop the UML Application Schema(s). Other UML tools or special templates in off-the-shelf editors may also be used but are likely to have minor differences in UML notations which will need to be adjusted or explained in the Product Specification.

XML data including Feature Catalogues and metadata is easier to view in open-source or COTS XML software rather than ordinary text editors.

#### **B-7.6.6 Identification of models**

The identification of each Application Schema must include a name and a version. If there is only one Application Schema in the Product Specification, this identification is implicit in the name and version of the Product Specification. Product Specifications with more than one Application Schema must identify each, potentially by associating it with a scope.

### **B-8 Data Classification and Encoding Guide**

The Data Classification and Encoding Guide (DCEG) provides information on how the data is to be captured. This should be as detailed and specific as necessary. The DCEG is used primarily by cartographers, editors, and data encoders, rather than application developers or OEMs; and should be written from that perspective.

The DCEG includes the collection criteria for mapping real world objects to the conceptual objects of the dataset.

Any organization performing data capture for the data product defined by the Product Specification must provide references to any more detailed encoding guide used in addition to that indicated in the Product Specification for the capturing process.

### **B-9 IHO GI Registry**

Procedures for registration are explained in S-99. It is recommended that at least one member of the Project Team or Working Group register as a Submitting Organization. Submitting Organizations propose changes and additions to the contents of Registers. Submitting Organizations will normally represent a recognized body or stakeholder group (such as from government, industry, academia and relevant user groups). Registered Submitting Organizations may submit proposals for consideration under any Domain in a Register. Stakeholders and any other interested parties who do not wish to enrol should submit proposals through an existing Submitting Organization.

To harmonize with other Product Specifications, developers should propose extensions to existing Registry items where possible. For example, propose generalization or specialization of an existing element; or additional values in an enumeration or codelist type. Restrictions of existing types can become new sub-types rather than changes to an already defined type; or it may suffice to define a constraint in the Product Specification.

### **B-10 Feature Catalogue**

The Feature Catalogue is an XML document which conforms to the S-100 XML Feature Catalogue Schema. Note, for Imagery and Gridded Data, a coverage is a feature type; and a Product Specification Feature Catalogue should define the attributes, coverage feature (with spatial primitive type 'coverage').

Feature Catalogues should be documented by a text-based documentation of their contents, which should also be reviewed by the Project Team and responsible Working Group. For review purposes, this text-based documentation should be generated from the XML Feature Catalogue. The resultant text can be in Word or PDF; or another format preferred by the team.

## **B-11 Data transfer modes and packaging**

Define mode of delivery – exchange set, message, or service. A Product Specification may also specify more than one delivery mode.

This needs to be done before metadata is specified because some metadata elements as well as the treatment of metadata (for example separate versus embedded) depend on delivery mechanisms, constraints, and protocols.

Details of packaging and transfer content can be finalized in a later stage (see clause B-18).

## **B-12 Metadata**

The minimum metadata requirements are set forth in Part 4 of S-100 (Appendix 4a-D for vector data, Parts 4b and 8 for coverage data). Product Specification developers should consider whether the metadata elements listed in S-100 are relevant to the data product and which of them are appropriate for its allowed packaging and delivery methods. For relevant elements, define appropriate values and restrictions if necessary for the metadata elements listed in S-100 Appendix 4a-D or Parts 4b and 8. Developers should note that Part 4b is quite skeletal in S-100 Edition 4 and the development team will need to use the underlying ISO standards and ISO metadata schemas.

### **B-12.1 Metadata for exchange set products**

This section describes metadata for products that are delivered in exchange sets.

#### **B-12.1.1 Generic metadata model**

S-100 provides for S-100 discovery metadata for exchange sets to be encoded in the Exchange Catalogue. Figure B-12-1 below depicts an example of relevant cases. The latest approved S-100 discovery metadata model is located within the latest edition of S-100 Part 4.

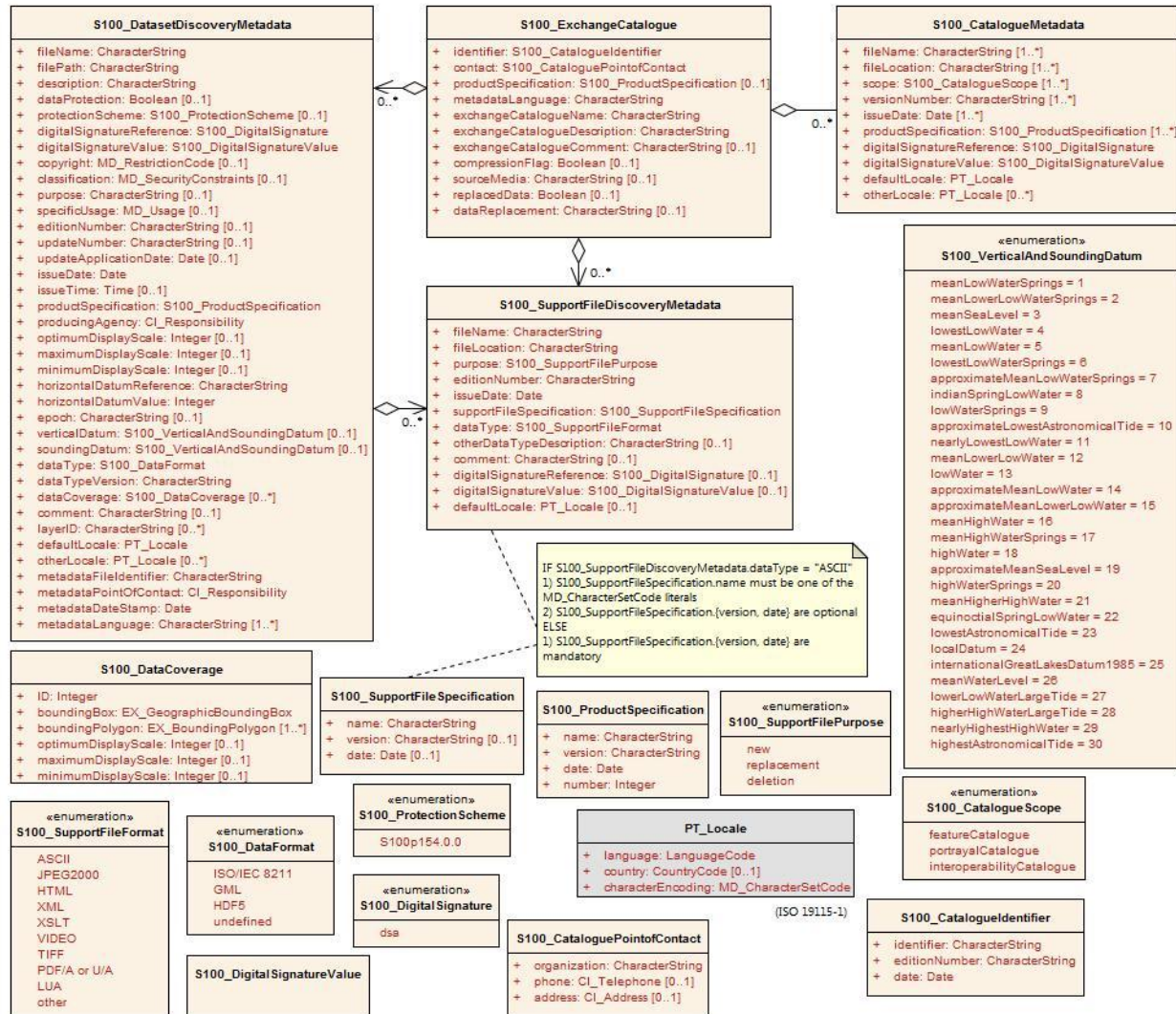


Figure B-6-3 – Exchange Catalogue and discovery metadata classes

If additional metadata elements are required they should be documented in the Product Specification Metadata section and extensions to the generic metadata schemas developed using the standard ISO extension mechanism.

Prior to the creation of extended metadata, a careful review of the existing metadata within ISO 19115-1 [and the S-100 generic metadata model] must be performed to confirm that suitable metadata does not already exist. If suitable metadata exist within ISO 19115-1 [or the generic S-100 metadata model], then it must be used. (S-100 Part 4a, Appendix 4a-E, extended to include S-100 generic metadata.)

IHO metadata XML schemas for Exchange Catalogues and discovery metadata have been developed and are available at the IHO software distribution site (<https://github.com/IHO-S100WG>) and will be served through the IHO GI Registry (<http://registry.iho.int/>) or another IHO site ultimately.

An Exchange Catalogue Builder to be provided by IHO is under development. Exchange Catalogues can also be prepared using off-the-shelf commercial and open-source XML editing and authoring tools. Generic

tools for ISO metadata can be used for ISO metadata files for each dataset (in S-100 Edition 4 and S-100 Edition 3, the ISO metadata files are separate from, and referenced by, the Exchange Catalogue XML files).

### **B-12.1.2 Use of the IHO S-100 generic metadata model and schemas in Exchange Catalogues**

If the generic S-100 Exchange Catalogue format and XML schemas cannot be used as is, either with or without restrictions, then Product Specifications must derive their metadata models from the generic model in S-100 Part 4a (supplemented with Part 4b, for gridded data). Derivations may be restrictions, extensions, or a combination of both. Derivations must conform to the rules in S-100 Part 4a, Appendix 4a-E. Custom UML diagrams similar to Figure 4a-D-4 in S-100 are required for Product Specifications that add classes or attributes, and recommended if Product Specifications omit optional classes/attributes or restrict allowed values in enumerations.

S-100 Part 4a specifies which of the S-100 metadata classes and attributes are mandatory and which are optional. Product Specifications may omit optional S-100 metadata classes or attributes as appropriate.

Restrictions on S-100 classes and attributes must be expressed as constraints. Restrictions include making an optional attribute mandatory or using a subset of attribute values. Restrictions do not require new metadata classes either.

The model can be extended with product-specific classes derived from the classes in the generic model. Derived classes can define additional attributes.

In order to implement a product-specific metadata model, the S-100 XML schemas that encode generic S-100 metadata can be supplemented with:

- Executable constraint checks to apply product-specific restrictions, in Schematron<sup>3</sup> or another language.
- Product-specific schemas that import the generic XML schemas and extend the generic XML types, to capture extensions.

This method means no change to the generic Exchange Catalogue XML schema or XSD files is needed. Instead product-specific customization can be implemented by adding supplementary files to the S-100 generic implementation.

Table B-12-1 below describes how Product Specifications can describe their metadata models. The actions are elaborations of the allowed extensions listed in S-100 Part 4a, Appendix 4a-E. The UML diagram would be based on S-100 Figure 4a-D-4.

Product Specifications must define their metadata models by either reusing the S-100 generic metadata model or extending the generic model in conformance with the rules in S-100, Appendix 4a-E.

ISO 19115-1 and S-100 Part 4a-E state: Prior to the creation of extended metadata, a careful review of the existing metadata within ISO 19115-1 must be performed to confirm that suitable metadata does not already exist. If suitable metadata exist within ISO 19115-1, then it must be used.

Unnecessary specializations of the generic S-100 metadata classes should therefore be avoided.

---

<sup>3</sup> Schematron rules can be checked using off-the-shelf software like XML editors, but implementations can implement the constraints in any suitable language for production tools or applications which cannot integrate Schematron validation or XSLT in their workflows.

| Action  | UML diagram  | Documentation table  | XML implementation   |
|---|--|--|--|
| Omit an optional class or attribute   | Suppress its display using diagramming tool functionality.   | Omit the corresponding table or row.   | Rule to check that the XML element <u>is not</u> present.  |
| Make an optional class or attribute mandatory   | No change. The multiplicity will still be 0.. but a diagram note may be added stating that it is mandatory in this product.                        | Multiplicity column should have the mandatory multiplicity and the Remark column a remark stating it is mandatory in this product.                     | Rule to check that the XML element <u>is</u> present.  |
| Restrict multiplicity   | As for the previous row, with appropriate modifications. This is the general case of making an optional attribute mandatory.                       |  |  |
| Limit enumeration values to a subset  | Suppress display of excluded values using diagramming tool functionality.  | Omit the corresponding row.  | Rule to check that the value <u>is not</u> used.   |
| Restrict the value of a text, numeric, CharacterString attribute  | No change required.  | Specify the value in the Remark column.  | Rule to check that the value <u>is</u> as specified.   |
| Add a new metadata class or attribute   | Specialize the appropriate S-100 metadata class and add the new class or attribute.  | Add a new documentation table, or extend the table describing the original class. Inherited attributes should be distinguished from direct attributes. | Additional XSD that imports the generic metadata XML schema and extends its types.   |
| Add a new metadata package in the Exchange Catalogue.<br>(Note: Whether the new package must be in the Exchange Catalogue or the ISO metadata depends on product-specific considerations like its expected use in applications) | Add the class or classes for the new package and aggregate to a specialization of the Exchange Catalogue container class (S100_ExchangeCatalogue). | New tables for the new classes and extended catalogue container class.   | As specified in ISO 19115-3 and S-100 clause 4a-5.6.5). If required in the Exchange Catalogue, additional XSD that imports the generic metadata XSD and extends its types. |
| Add a new metadata package only in ISO metadata   | Diagram(s) extending S-100 Figure 4a-1 and specifying new classes and attributes.  | New tables for the new classes.  | Describe the extensions in the ISO 19115-3 metadata extension format and extend the ISO metadata XSDs with new types.  |

| Action   | UML diagram  | Documentation table   | XML implementation  |
|--|--|---|---|
| Add values to an enumeration<br>(Consider proposing its addition to S-100 generic metadata via an S-100 maintenance proposal instead of a product-specific extension)  | Specialize the relevant S-100 metadata class and override the affected attribute.  | Describe the override and new data type in the appropriate tables.  | XML schema override in a derived XSD. Ignore any warnings generated by generic constraint-checking rules. Define new constraint-checking rules if necessary.  |
| Create new metadata codelist elements<br>(expand a codelist)   | Expand codelist in diagram(s). (Some codelists like languages may be too long for a UML diagram.)                                      | Add codelist values to codelist table (if any), or document the new values using another method.              | Add new value to appropriate codelist and GML dictionary files.   |
| Create a new metadata codelist to replace the domain of an existing metadata element that has "free text" listed as its domain value (Note: Simply restricting the character strings as described in an earlier row is a simpler solution) | Specialize the relevant S-100 metadata class and override the affected attribute. Add a codelist UML element documenting the codelist. | Document the specialization and new codelist in appropriate tables.   | XML schema override in a derived XSD. Ignore any warnings generated by generic constraint-checking rules. Define new constraint-checking rules if necessary. Add new codelist to appropriate codelist and GML dictionary files. |
| No value for a mandatory attribute   | None.  | Specify that the mandatory attribute must be nilled or assigned one of the allowed values as a fixed default. | Rule to check the value.  |

*Table B-12-1 – Specifying metadata in Exchange Catalogues for individual Product Specifications*

Product Specifications should not clone and rename classes, attributes, and enumerations defined in the generic model, but instead extend the generic model as needed. This requirement helps keep product-specific and generic S-100 metadata harmonized as both S-100 and the Product Specifications evolve over successive versions. Cloning and renaming classes and attributes will result in different branches of metadata. The divergence of metadata branches will increase as both S-100 and Product Specifications evolve over successive versions. One consequence will be a need for implementations to have independent product-specific metadata processing modules. This situation would be undesirable for products that are processed by the same application (such as generic viewers, ECDIS, or ECS).

A typical Product Specification Exchange Catalogue UML diagram is reproduced in Figure B-12-2 below. Compare to Figure B-12-1, noting the omission of attributes for vertical and sounding datum from dataset discovery metadata and restrictions of support file format and data format enumerations.

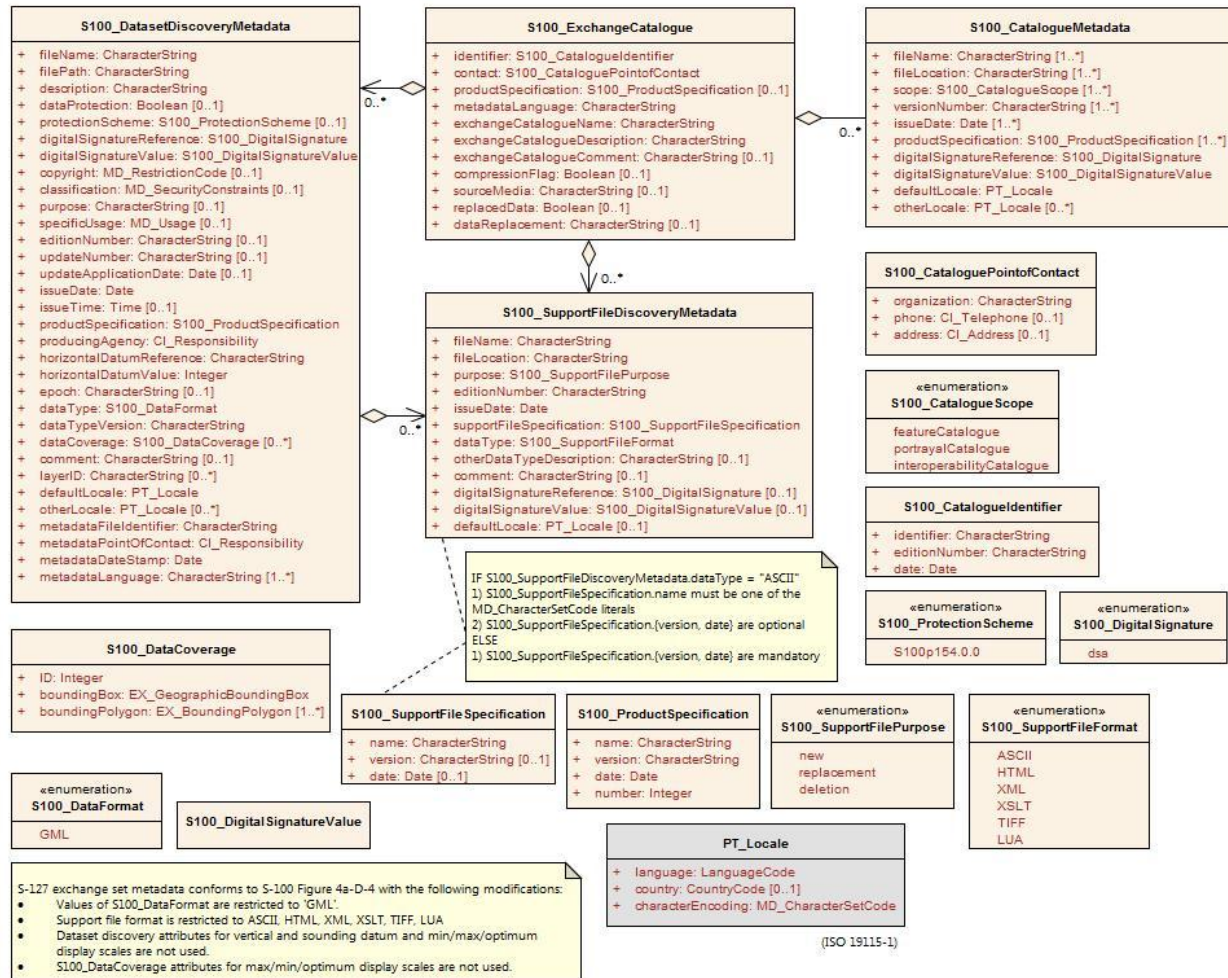


Figure B-12-2 – Example Exchange Catalogue metadata classes for S-127 (Marine Traffic Management)

### B-12.1.3 Extensions in ISO metadata files

S-100 Part 4a, clause 4a-5.6.5 specifies how extensions must be documented. This specification is derived from the ISO Specification.

### B-12.1.4 Data protection, authentication and encryption

The purpose of data protection in S-100 is threefold:

- 1) Piracy protection: To prevent unauthorized use of data by encrypting the product information.
- 2) Selective access: To restrict access to only the products for which a customer has acquired a license.
- 3) Authentication: To provide assurance that the products were actually created and distributed by the producers and distributors who are identified as such in the product package or datasets.

Piracy protection and selective access are achieved by encrypting the products and providing data permits to decrypt them. Data permits have an expiration date to enable access to the products for a licensed period. Authentication is provided by means of digital signatures applied to the product files. Selective access to individual products is supported by providing users with a licensed set of data permits. The license



is created using a unique hardware identifier of the target system and is unique to each Data Client. Consequently, licenses cannot be exchanged between individual Data Clients.

#### **B-12.1.4.1 Application of protection measures**

Product Specifications should specify whether datasets must, must not, or may be encrypted using the S-100 recommended security scheme, which utilizes the *dataProtection* attribute of **S100\_DatasetDiscoveryMetadata**. If this Boolean attribute is set to TRUE the co-attribute *protectionScheme* must also be assigned a value from the enumeration of security schemes (**S100\_ProtectionScheme**). S-100 at present defines only one recommended security scheme, which is described in Part 15.

#### **B-12.1.4.2 Digital signatures**

S-100 Part 4a mandates digital signatures for datasets and catalogues in exchange sets. For support files, digital signatures are allowed but optional. Product Specification authors should therefore specify which support files must, must not, or may be signed. Signing datasets is independent of whether the recommended security scheme is applied to the dataset.

The digital signature method is encoded in attribute *digitalSignatureReference*. There is also an attribute *digitalSignature* for encoding the digital signature itself. These attributes are present in dataset discovery metadata, catalogue metadata and support file discovery metadata (classes **S100\_DatasetDiscoveryMetadata**, **S100\_ExchangeCatalogue**, and **S100\_SupportFileDiscoveryMetadata**). The structure of the digital signature is specified in S-100 Part 15. S-100 Part 4a is quite permissive as to the format, allowing either well-formed XML (one or more XML elements from an unspecified schema) or a character string (for example base64 encoded signature).

#### **B-12.1.4.3 Compression and archive format**

Compression of data products as used in S-100 Parts 4a and 15 includes specification of the archive format as well as the actual file compression method. In S-100 Edition 4.0.0 there is only one archive format (ZIP) and only one allowed compression method (DEFLATE). Compression requires packing into an archive. Product Specification authors must specify whether an exchange set must, must not, or may be compressed by specifying appropriate constraints on the *compressionFlag* metadata attribute in Exchange Catalogues. Further information about archive format and compression method is provided in S-100 Part 15, clause 15-5. S-100 Edition 4.0.0 provides for Exchange Catalogues to have only one instance each of the *compressionFlag* attribute, which therefore applies to all files in the exchange set (that is, after compression, there will be only one ZIP archive which contains all data files, support files, and catalogues in the exchange set, with the DEFLATE compression method applied to all<sup>4</sup>).

Product Specification authors should note that an exchange set may contain other exchange sets. Each included exchange set can be treated as an individual item for compression purposes (that is, packed into a ZIP archive or not), but they will all be packed into the archive of the overall container exchange set, either as a folder hierarchy or single-file ZIP archives, depending on whether they are individually compressed.

The encryption and digital signature features of ZIP are not used.

---

<sup>4</sup> Some Zip implementations may check whether the compression method actually produces a smaller file, and if not, store the original instead. It may be possible to force compression even if the file size is not reduced.

### B-12.1.5 Metadata for imagery and gridded data

Exchange set metadata for Product Specifications dealing with imagery and gridded information is the same as for ordinary vector datasets (clauses B-12.1.1 – B-12.1.4), extended with additional metadata elements and attributes specific to imagery or gridded data. S-100 Part 4b describes the additional metadata elements, which are defined in detail in ISO 19115-2.

The Exchange Catalogue for such products will be as for vector data products. As for vector data, the Product Specification team may define additional product-specific metadata elements and attributes in conformance with S-100 Appendix 4a-E. ISO-defined elements and attributes should be in the ISO metadata file.

### B-12.1.6 Embedded metadata

Certain metadata may be embedded in dataset headers (ISO 8211 and GML formats – S-100 Parts 10a and 10b) or defined attributes and groups (HDF5 format – Part 10c). While specifying embedded metadata is up to individual Product Specifications, this guideline recommends that only metadata considered essential to identifying and reading the dataset should be embedded. Examples are dataset name, persistent global identifier and MRN (if any); as well as bounding box, number of objects of each type and CRS identification. Gridded data products may encode spatial representation information such as grid spacing and grid bounding box as metadata for individual features. Gridded data products may also encode structural metadata such as a code indicating the type of grid (regular, irregular, etc).

## B-12.2 Metadata for services

This section describes metadata for products that are delivered as services.

### B-12.2.1 Generic metadata for services

Generic metadata for services is a work in progress. S-100 Part 4a defines a metadata model for services which conforms to the ISO 19115-1/2 model. S-100 Part 14 defines service metadata differently (**S100\_ServiceMetaData** – clause 14-8.1.1). It is left to Product Specifications to determine if servers need to supply service identification metadata in the sense of the underlying ISO standards (reproduced in S-100 Figure 4a-A-2).

Product Specifications needing to define service metadata (identifying and describing the service itself) should use or extend the **SV\_ServiceIdentification** class in S-100 Figure 4a-2. Any extensions or restrictions should conform to the ISO rules as described in S-100 Part 4a, Appendix 4a-E.

### B-12.2.2 Use of the IHO S-100 metadata model and schemas

S-100 Part 14, clause 14-9.1.3 specifies a *GetMetaData()* operation with *CharacterString* return parameter *exchangeSet* defined as “The exchange set describing the datasets”. It is not clear at this time whether this is the metadata defined in S-100 Part 14, clause 14-8 or the Exchange Catalogue of S-100 Part 4a, Figure 4a-D-4. This can be considered **payload metadata** (metadata describing the transferred information) to distinguish it from metadata that describes the service itself.

If Product Specifications need to define a subset and/or extension of the S-100 Exchange Catalogue to describe payload metadata, they should use the same methods specified in S-100 Part 4a/4b and clause B-12.1 of this Guidance. Note that these sections require conformance to the ISO rules described in S-100 Part 4a, Appendix 4a-E.

### B-12.2.3 Data protection, authentication and encryption

Many of the data protection, authentication and encryption provisions of S-100 Part 15 should also apply to services, but details and implementation constraints will differ. For example, a data product may create its transfer package using a format other than ZIP; and the connection protocol may need to be secure (for example HTTPS rather than HTTP). Product Specification teams which need to use security schemes in service transfer modes are encouraged to use or adapt S-100 Part 15 and report on their experiences.

### B-12.2.4 Embedded metadata

The considerations are similar as for embedded metadata described in clause B-12.1.6. Additional considerations may arise from the transfer mode; for example, giving greater importance to minimizing data volume. Product Specification teams are encouraged to use or adapt the embedded metadata principles in clause B-12.1.6 and the methods and constraints of the data format which is used (ISO 8211, GML, HDF5, or other format) and report on their experiences.

## B-13 Define data encoding format

### B-13.1 Selection of encoding format

The encoding format should be selected based on the type of product and other requirements, including production and processing. The characteristics of the three standard data formats included in S-100 Edition 4.0.0 are summarized in Table B-13-1 below for convenience.

|                                   | ISO 8211                                      | GML   | HDF5   |
|-----------------------------------|---|---|--|
| Type of product to which suited   | Nautical charts and feature-heavy vector data | Nautical publications and information-heavy vector data; discrete weather information; small datasets such as marine safety information; data delivered via messages and web services | Coverage-based data                          |
| Generic data format               | Yes   | Yes   | Yes  |
| Data production complexity        | Requires custom tools                         | Can be produced with a range of tools from text editors to custom apps and database SQL queries   | Custom apps that use off-the-shelf libraries |
| Processing complexity             | High  | Low   | High   |
| Supporting off-the-shelf software | Yes   | Off-the-shelf viewers and server software; can be viewed with ordinary text editors   | Off-the-shelf viewer                         |
| Data volumes                      | Lower   | Higher  | Lower  |

|   | ISO 8211          | GML   | HDF5   |
|---|-------------------|---|--|
| Type of data  | Vector            | Vector; coverage schemas are defined in the GML specification but not used in S-100   | Gridded  |
| Supporting artifacts needed in Product Specifications | Feature Catalogue | XML schemas for data validation; datasets can be processed by apps without XML schemas; self-documented format (tags indicate objects and attributes); Feature Catalogue optional | Embedded object and attribute tags; Feature Catalogue optional |

Table B-13-1 – Characteristics of S-100 standard encoding formats

While other formats than the three standard encodings are possible, the use of a non-standard format has the following implications:

- Loss of genericity, requiring special purpose development by implementers or conversion to a standard format with impacts on performance.
- Potential loss of compatibility with interoperability.

The question of appropriate formats for transactional, web-service, or message-based information has not been formally addressed in depth at this time, but Product Specifications needing such delivery modes should endeavour to use one of the standard formats in order to minimize implementation complexities.

### B-13.2 Data format definition artefacts for GML

Selection of the GML format will require definition of XSD files encapsulating the S-100 Application Schema and conforming to the GML specification (ISO 19136 and S-100 Part 10b). There must be a 1/1 mapping from the Application Schema to XML schema constructs in the GML XSD files that implement the model as GML Application Schemas.

- Feature and information classes in the UML domain model map to XML complex types and XML elements. The class (camel case) name is the same as the element name (XML tag excluding any namespace).
- Simple attributes in the UML domain model map to XML elements contained in the feature and information complex types mentioned above. The names (XML tags, excluding any namespace) are the same as the (camel case) names of the UML attributes.
- Complex attributes in the UML domain model map to XML elements contained in the feature and information complex types mentioned in the first bullet. The names (XML tags, excluding any namespace) are the same as the (camel case) names of the UML attributes. The types of these XML elements are XML complex types with the same (camel case) name, either with or without “Type” suffixed.
- Enumerations map to XML enumerations of the same (camel case) name.
- Codelists map to XML union types of the same (camel case) name; one member of the union type is an anonymous enumeration of the listed values and another member is a string restricted to the pattern “other: ...” as specified in S-100 Part 3, clause 3-6.7.

- Associations are encoded as elements contained in the feature and information complex types, named according to the role indicated in the UML domain model and with a GML attribute that contains the gml:id of the referenced object.

Schema developers should note that conformance to the GML specification requires conformance to the rules set forth in the GML specification (ISO 19136 / OGC 07-036), not merely XML-validation against the GML schemas.

S-100 Part 10b describes the S-100 GML profile. The XSD files for the S-100 GML profile are available at the IHO S-100 distribution site (<https://github.com/IHO-S100WG>). Previously defined GML schemas for other Product Specifications (for example S-122, S-123) will provide useful guidance and will be made available through the IHO Product Specification distribution site or GitHub distribution site.

Guidance for generic processing for GML datasets is included in S-100 Edition 4.0.0.

## B-14 Portrayal elements and rules

Portrayal Catalogues are necessary only for data products that are intended to be displayed graphically (as opposed to text or other processing).

Often there will be only one intended application domain – generally ECDIS – and a limited number of closely related usage scenarios (that is, route monitoring/planning and voyage planning) utilizing one Portrayal Catalogue. Some data products may not be intended for display, in which case no Portrayal Catalogue is needed. Other data products may be intended for more than one application domain, in which case the possible need for different Portrayal Catalogues should be considered. Different usage scenarios may also benefit from different Portrayal Catalogues.

S-100 defines an XSLT-based portrayal mechanism in Part 9. Edition 4.0.0 adds the LUA scripting language as a second portrayal mechanism and defines a standard mechanism for including scripting support in S-100 based products. Developers/OEMs may prefer LUA due to promises of performance and similarity to common programming languages.

The portrayal section of a Product Specification should include:

- Pictorial representations of symbols and colours, accompanied by symbol specifications (the latter preferably in the form of machine-processable files as well as formal specifications in the text of the portrayal section).
- Specifications and recommendations for the use of symbols by implementers; for example, calculating orientation from attribute values and the use of thinning to reduce crowding of displays at small display scales, display of text accompanying symbols, masking of boundaries, etc.
- Products that need to be included in ECDIS must have day, night, and dusk palettes defined.

## B-15 Registration of portrayal elements

Portrayal elements must be registered in the Portrayal Register of the IHO GI Registry. Elements that need to be registered are listed below. A portrayal element needs to be registered only if it is new<sup>5</sup> to the Register – elements that are already in the Register need not be registered again.

- New symbols, fonts, line styles, area fills, and pixmaps.
- New display modes.
- New viewing group layers.
- Viewing groups.
- New display planes.
- New colour tokens.
- New colour profiles used by portrayal of this product. If there are multiple palettes (for example day/night/dusk), the colour profile file(s) must define all palettes.
- New CSS files encoding the colour tokens and RGB values used by all symbols in the portrayal of this product.
- New context parameters.
- New display priorities.

## B-16 Portrayal Catalogue

An IHO Portrayal Catalogue Builder is available to Product Specification developers in the Help and Guidance/Repository section of the IHO GI Registry.

## B-17 Reference systems

The preferred coordinate reference system is EPSG 4326 which is based on the WGS 84 horizontal datum.

Horizontal datum will normally be referenced by giving its code in the EPSG Register. If the coordinate reference system is not one of the coordinate reference systems in the EPSG Register, a datum may be specified in a support file as described in S-100 Part 6, should there be a use case in the scope of the data product.

A set of vertical datums is listed in S-100 Appendix 4a-D (S100\_VerticalAndSoundingDatum). Product Specification developers are encouraged to adopt the S-101 ENC datum (either standard or local S-101 ENC datum) as a common vertical reference datum if possible. If a need for an additional datum is identified, it should be proposed as a revision to S-100.

S-100 includes 'local datum' as an allowed value for datum attributes, but this is of limited utility even within a data product. (S-101 cites an example of use of local datum in a non-tidal basin, where depths may refer to a sounding datum different from that in open waters. If this area is navigable at the maximum display scale of the ENC data, the value of this datum must be encoded using attribute vertical datum = 24

---

<sup>5</sup> This clearly introduces a dependency between different data products. The Portrayal Register design and business processes to manage such dependencies need to be worked out. This section may have to be revised after that is done.

(local datum), in a meta-feature co-incident with the area covered by the dock.) Data conversion to/from unspecified local datums would be problematic.

## B-18 Data product delivery

### B-18.1 Delivery content and structure

Define the content and structure of delivery packages: Exchange sets, messages, or web services.

#### B-18.1.1 Exchange sets

Exchange set structure should be defined, either by using the structural diagram from S-100 Appendix 4a-D (reproduced in Figure B-18-1 below) as is; or restricting the allowed components; or defining extensions of the individual components. If there is internal structure in the exchange sets (for example folders and sub-folders), determine the required layout and naming conventions. Determine how the exchange set as a whole is packaged (for example ordinary folders, zip file, etc.).

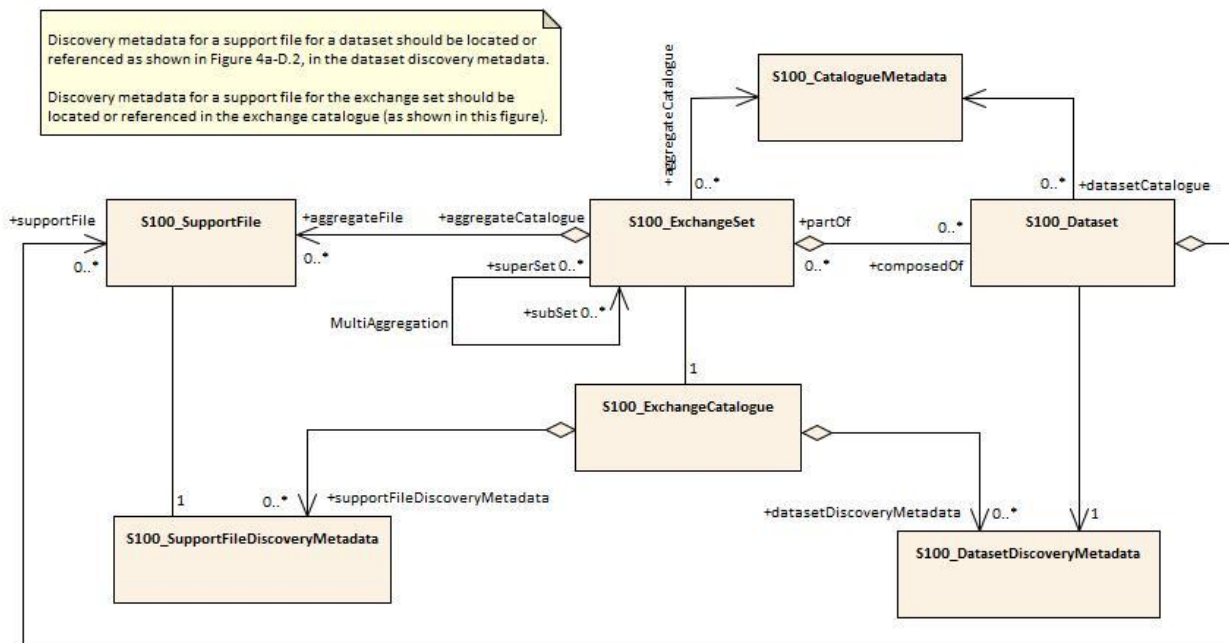


Figure B-18-1 – Prototype exchange set structure

Determine what naming conventions, if any, must apply to individual components of the exchange set - dataset files, exchange sets and support files. Naming conventions are generally not needed for message-based or service modes of data delivery, but a unique identifier will generally be needed for each message or service transaction.

#### B-18.1.2 Services

S-100 Part 14 describes three types of communication:

- Session oriented communication (S-100 Part 14, clause 14-4) - point-to-point connections between client and server described by an interaction model that describes the life span of a session (initiation, maintenance and termination of the session).
- Session-less interactive communication (S-100 Part 14, clause 14-5) - an encapsulation of all relevant information within a request, based on which the server formulates an appropriate response.
- Message streams (S-100 Part 14, clause 14-6) - a unidirectional flow of messages containing well-defined sets of data.

For message and web service modes of delivery, specify the container format and packaging; and specify the transfer protocol (for example REST (Representational state transfer), SOAP (Simple Object Access Protocol)) and packaging (for example WSDL (Web Services Delivery Language), WFS (Web Feature Service)). S-100 Part 14 currently permits three service technologies: SOAP, REST and CORBA.

[Developing specifications for online services will be included in a later version of this document, pending more experience with S-100 Part 14.]

## **B-18.2 Dataset updates**

### **B-18.2.1 General considerations for updates**

Define the conditions and mechanisms for data updates, specifically:

- Update cycles – how frequently data must be updated; whether updates are issued on a regular cycle, as-needed, or a combination of both;
- How long each dataset is valid and how validity periods will be indicated;
- Mechanisms for cancellation, replacement and reissue of datasets;
- Metadata for updates;
- Types of updates – whether the data product requires incremental, whole-dataset replacement, irregular or cumulative updates; and
- Criteria for determining when datasets must be replaced by new datasets, superseded, reissued, updated and cancelled.

### **B-18.2.2 Format-specific update considerations**

ISO 8211: S-100 Part 10a describes update mechanisms for ISO 8211 datasets. Each Product Specification must define the structure of update datasets.

GML: Replacement of whole objects is the recommended method, but there are XML specifications that allow update of individual attributes. (Details will be included in a later version of this document, pending further discussions.)

HDF5: Feature (coverage) can be updated in its entirety or in part (the update can be a sub-grid).

## **B-18.3 Supporting information**

Describe how any auxiliary content is delivered either with or as an adjunct to data. S-100 provides for 'support files' to be included in exchange sets. Support files can be graphic or text information files



referenced by dataset objects, or other files such as dictionaries and catalogues (including Feature or Portrayal Catalogues). Define allowed file formats and naming conventions for support files.

Note that since Feature and Portrayal Catalogues are shared by all datasets conforming to a specific version of a Product Specification, it will generally be more efficient to deliver Feature and Portrayal Catalogues once rather than with every exchange set. Possible methods of such special deliveries have not been standardized yet and are left to Product Specification developers, but may include special exchange sets distributed through the usual channels or a centralized means such as publication on a web server. If a central distribution mechanism is adopted consideration must be given to the possible needs of end users who have only infrequent or no access to the distribution hub (for example low-bandwidth or no Internet access).

## **B-19 Validation checks and data quality**

At least two types of validation checks are needed:

- Dataset validation checks for individual datasets. These checks operate on individual objects in datasets and on individual datasets as a whole. They should check the integrity of individual objects in the dataset (spatial, feature and information types); associations between objects in the dataset; any embedded metadata or header information in the dataset; and support files referenced in the dataset.
- Package validation checks for verifying the structure and content of packages (for example exchange sets) and accompanying metadata.

A common set of validation checks is under development (see S-97 Part C). A recommended set of data quality measures has been developed based on dataset statistics derived from the validation checks. The recommended measures are also described in S-97 Part C.

### **B-19.1 Validation checks for datasets**

Validation checks for datasets should cover:

- 1) Completeness; including population of attributes and presence of required information, complex attributes without sub-attributes, etc.
- 2) Logical consistency; for example, missing association targets.
- 3) Spatial consistency; for example, topological sanity checks for non-crossing external boundaries, excessive vertex density in lines, etc.
- 4) Positional accuracy; for example, closeness of reported coordinate values to accepted or known absolute or relative coordinate values.
- 5) Temporal accuracy; for example, closeness of reported time measurements to accepted or known values accepted as or known to be true, correctness of the order of events, or validity of data with respect to time.
- 6) Thematic accuracy; such as attribute values that are consistent with any other related attributes and within allowed ranges or sets.
- 7) References to support files.
- 8) Other requirements specific to the product; for example, encryption, signatures, etc.

Some of these issues are addressed in the common set of validation checks. A recommended common set of validation checks is described in S-97 Part C. Product Specification developers should supplement the common set with such additional checks as are appropriate for the particular products.

### **B-19.2 Validation checks for packages**

Validation checks for packages should cover:

- 1) Package completeness – whether all required components are included, including datasets, support files, metadata and appropriate catalogues (for example exchange set catalogues, Feature Catalogues and Portrayal Catalogues). Note that the Product Specification must indicate which catalogues are appropriate to the delivery method; for example, message-based delivery methods may not include catalogues in the delivery packages.
- 2) Package container format and structure – whether the package is in the approved container format (for example ISO 8211, TIFF, etc) and whether appropriate encryption and signatures have been applied at the container level. Examples of package validation checks are:
  - Assuming the Product Specification specified delivery as zip files, is the container a zip file of the appropriate type?
  - If the package is arranged in a directory (folder) structure, are the structure and names of directories (folders) as required in the Product Specification?

Package validation checks are required to validate delivery packages, but are expected to be out of the scope of S-97 Part C and Product Specification teams may have to specify their own. The tests for exchange set and service delivery modes will obviously be different, but the matter of validating the delivery package or stream should be addressed for all delivery modes covered in the Product Specification.

### **B-19.3 Common validation checks**

Given that some features, information types and Application Schema constructs are used in multiple products, there will be validation checks in common with existing Product Specifications and any such related Product Specifications should be consulted for validation checks. Spatial consistency checks in particular, as well as consistency checks related to meta-features, can be expected to be in common with several Product Specifications.

Spatial operations used in validation checks must be the operations defined in IHO ENC Validation Checks (S-58 Edition 6.0.0 or its successor).

### **B-19.4 Validation checks for base versus update datasets**

If the Product Specification defines an update dataset format, the validation checks developed for new (base) datasets should be reviewed for their applicability to update dataset formats.

## **B-20 Preparing for interoperability**

Product Specification developers should carry out the activities described below if the data product is a candidate for inclusion in an Interoperability Catalogue. Further details on these activities will be provided in a future edition of S-100 and in IHO Publication S-98 (in preparation).

- Determine which, if any, product groups in Interoperability Catalogues are supplemented or enhanced by the data product.

- Determine whether, and how, the IHO Interoperability Catalogue will be affected by the new product; including updates to display priorities, interleaving, predefined combinations and other interoperability rules and operations.
- Revise Portrayal Catalogues upon recommendation by the IHO Interoperability Team.
- Compare the Application Schema to related Product Specifications and advise the developers of the Interoperability Catalogue on their similarity to features in other Product Specifications. This comparison should include feature concept similarity, attribute bindings, attribute value domains, potential discrepancies in feature geometries, etc.
- Examine the Specification scope and factors affecting the quality of information which may also be present in other products; for example, whether information is for context or background in one product but a significant part of the other. The results of this analysis should be communicated to Interoperability Catalogue developers.
- Consider interactions due to cartographic objects; for example, text placement issues in interoperability mode. Product Specification teams should consult with OEMs on how to mitigate such issues.
- Advise Interoperability Catalogue developers on whether features are supplementary information for features in another product. Such features may be combined as part of interoperability.
- Attempt to harmonize effects of maximum and minimum display scales used by datasets in order to avoid situations where one data product is within display scale but a complementary product is not.

## **B-21 Sample data / test datasets**

Test data should be created in sufficient quantity to validate the main characteristics of the Application Schema. Specifically, the first test dataset should contain:

- At least one instance of each kind of feature and information type.
- A representative set of feature and information associations, preferably at least one instance of each named association. (It is not necessary to create an instance for each and every pair of object classes which may be linked by an association.)
- At least one instance of each meta-feature and data quality feature.

At least one update dataset should also be prepared, to validate the update dataset format and packaging.

Additional test datasets should test typical data volume, representative data capture problems and error cases.

If delivery is intended to be in the form of exchange sets, the test datasets should be packaged as complete exchange sets, including sample metadata files. Sample packages for other forms of delivery (transactions, messages, web services, etc) should be emulated as realistically as is practical at this stage; that is, setting up a web server, service broker, etc for web services should be done if doing so is practical but is not an essential requirement (it can and should be done as part of the testing stage).

## **B-22 Testing and feedback**

A formal test plan should be prepared, including test cases.

## **B-23 Work processes**

### **B-23.1 Registration and requesting an S- number**

Apply for a Product Specification number – this should ordinarily be assigned when the development project is approved during the initiation stage.

Registration of Product Specification artefacts in the IHO GI registry should be done in accordance with the procedures established in IHO publication S-99.

### **B-23.2 Project Teams**

Project Teams should involve domain specialists, information modelling specialists and representatives of OEM/developer communities.

### **B-23.3 Iterative refinement as a development process**

Development should plan for iterative refinement, with the following being reviewed at the indicated stages of development:

- Initial Application Schema. Reviews of subsequent revisions can be rolled into the reviews of the main Product Specification document;
- First drafts of the main document of the Product Specification and DCEG. Subsequent revisions should be reviewed as ready;
- Feature Catalogue, data format and sample datasets should be checked after each major revision to the Application Schema and Feature Catalogue;
- Portrayal Catalogue – first draft and significant revisions;
- Other artefacts or components, such as validation tests – when substantially completed; and after revisions due to changes to the Application Schema, Feature Catalogue, or data format;
- The Product Specification as a whole – after the complete package is ready.

Reviews of different components can be combined to fit the development schedule or Working Group meeting schedules. The Application Schema, main Product Specification and DCEG should be expected to undergo multiple reviews during different stages of development.

Reviews during the development stage should be requested from:

- Project Team – initial reviews;
- Technical group(s) sponsoring the Specification, as well as related technical Working Groups – after some stability has been achieved;
- Developers, implementers, and OEMs – formal reviews after initial reviews and stabilization in the Project Team and sponsoring technical Working Group. Note that individual implementer/OEM/developers should ideally be involved from the earliest stages of development if available.

Reviews should be conducted for completeness, correctness, ability to capture and express the domain, performance/efficiency; as well as conformance reviews for verifying compliance to S-100 and underlying Standards.

In addition, the IHO Data Quality Working Group should review the Product Specification for completeness.

Stakeholder reviews should be requested as the Specification matures, and should involve:

- Producers;
- Developers and OEMs.

Subsequent stages should involve users and user testing; and preparation of an impact study.

Test development and testing should commence upon the Feature Catalogue and data format achieving reasonable stability, presumably after one or two cycles of review.

Final assessment will be at the HSSC level for IHO; or equivalent for other organizations.

Pre-publication review will be conducted by the IHO or other publishing organization prior to publication, to check production issues.

#### **B-23.4 Maintenance of Product Specifications**

Clarifications, corrections, and revisions should be designated in accordance with the same criteria used for S-100 as described in S-100 Part 12, Clause 12-2 (Maintenance Procedures).

Product Specifications should undergo periodic review. A two-year review period is suggested for new Specifications, which may be increased to five years after the Specification reaches maturity.

Page intentionally left blank

## Part C – Data Quality

### C-1 Overview

S-100, the Universal Hydrographic Data Model, is a hydrographic geospatial data standard that can support a wide variety of hydrographic-related digital data sources; and is fully aligned with mainstream international geospatial standards, in particular the ISO 19000 series of geographic standards. This alignment enables easier integration of hydrographic data and applications into geospatial solutions. S-100 is inherently more flexible than S-57 and makes provision for such things as the use of imagery and gridded data types, enhanced metadata and multiple encoding formats. It also provides a more flexible and dynamic maintenance regime for features, attributes and portrayal via a dedicated online Registry. S-100 provides a framework of components that enables the building of standardized Product Specifications for the modeling of hydrographic data, thus providing true interoperability between different data standards and systems.

S-97 is a Guideline intended for developers and maintainers of Product Specifications that utilize the IHO framework standard S-100 (Universal Hydrographic Data Model).

### C-2 Introduction

Part C is intended to ensure that data quality aspects are addressed in an appropriate and harmonized way for all S-100 based Product Specifications.

This Data Quality Guideline can be used by HSSC Working Groups developing S-100 based Product Specifications. It provides 10 recommendations for appropriate data quality measures as deemed necessary to be used within S-100 based Product Specifications.

When drafting a Product Specification, the Data Quality Guideline will serve as a guidance document to verify if the appropriate Data Quality Elements have been included in the Product Specification. A Data Quality Element is a quantitative component documenting the quality of a dataset. The applicability of a Data Quality Element to a dataset depends on both the dataset's content and its Product Specification, the result being that all available Data Quality Elements may not be applicable to all datasets.

The place of data quality measures in dataset and exchange set metadata and the encoding of data quality in metadata is described in various ISO standards (ISO 19115, 19139, 19115-1/2/3, ISO 19157) and in S-100 Parts 4a–4c.

### C-3 References

- IG-D2.8.II.1 D2.8.II.1 INSPIRE Data Specification on Elevation – Technical Guidelines.
- ISO 8211 Specification for a data descriptive file for information interchange structure implementations. ISO/IEC 8211, 1994.
- ISO 19115 Geographic information – Metadata (2003). As amended by Amendment 1, 2006.
- ISO 19115-1 Geographic information – Metadata – Part 1 – Fundamentals. ISO 19115-1, 2014, as amended by Amendment 1, 2018.

|             |   |
|-------------|---|
| ISO 19115-2 | Geographic information – Metadata - Part 2 – Extensions for imagery and gridded data. ISO 19115-2, 2009.      |
| ISO 19115-3 | Geographic information – Metadata - XML schema implementation for fundamental concepts. ISO/TS 19115-3, 2016. |
| ISO 19139   | Geographic information – Metadata – XML schema implementation.  |
| ISO 19157   | Geographic information – Data Quality. ISO 19157, 2013, as amended by Amendment 1, 2018.                      |
| S-44        | IHO S-44 – Standards for Hydrographic Surveys.  |
| S-100       | IHO S-100 – Universal Hydrographic Data Model, Edition 4.0.0, December 2018.                                  |
| S-101       | IHO S-101 – Electronic Navigational Chart Product Specification, Edition 1.0.0, December 2018.                |
| S-102       | IHO S-102 – Bathymetric Surface Product Specification, Edition 1.0.0, April 2012.                             |
| S-121       | IHO S-121 – Maritime Limits and Boundaries, Edition 1.0.0, October 2019.                                      |

NOTE: In this document, “S-100” means S-100 Edition 4.0.0 unless a different edition is explicitly identified.

## C-4 Terms and abbreviations

### C-4.1 Terms

#### **data product specification**

a detailed description of a dataset or dataset series together with additional information that will enable it to be created, supplied to, and used by another party. A data product specification provides a description of hydrographic concepts and a specification for mapping the universe of discourse to a dataset. It may be used for production, sales, end-use or other purposes

#### **data quality element**

quantitative component documenting the quality of a dataset [ISO 19101:2002; S-100 4.0.0 Annex A]

NOTE: The applicability of a data quality element to a dataset depends on both the dataset’s content and its product specification, the result being that all data quality elements may not be applicable to all datasets

#### **data quality evaluation procedure**

the whole of operations used in applying and reporting quality evaluation methods and their results

#### **data quality measure**

an evaluation of a data quality sub-element

#### **data quality overview element**

the non-quantitative component documenting the quality of a dataset. Information about the purpose, usage, and lineage of a dataset is non-quantitative quality information

#### **data quality result**

a value or set of values resulting from applying a data quality measure or the outcome of evaluating the obtained value or set of values against a specified conformance quality level



**data quality scope**

the extent or characteristic(s) of the data for which quality information is reported

NOTE: The scope for a dataset can comprise a dataset series to which the dataset belongs, the dataset itself, or a smaller grouping of data located physically within the dataset sharing common characteristics. Common characteristics can be an identified feature type, feature attribute, or feature relationship; data collection criteria; original source; or a specified geographic or temporal extent.

**data quality sub-element**

a component of a data quality element describing a certain aspect of that data quality element

**Metadata ISO19115**

quality report in the format specified in the relevant ISO standards (ISO 19115-1/2/3 and ISO 19157)

**metaquality**

information about the reliability of data quality results [ISO 19157]

**result scope**

scope of the (data quality) result [ISO 19157, adapted]

NOTE: Result scope is a subset of the data quality scope.

**standalone quality report**

free text document providing fully detailed information about data quality evaluations, results and measures used [ISO 19157:2013]

**C-4.2 Abbreviations**

|       |   |
|-------|---|
| DQWG  | IHO Data Quality Working Group  |
| ECDIS | Electronic Chart Display and Information System   |
| ECS   | Electronic Chart System   |
| ENC   | Electronic Navigational Chart   |
| GI    | Geospatial Information (generally followed by “registry” or “register,” meaning either the IHO Geospatial Information Registry or a specific register in it). |
| GML   | Geography Markup Language   |
| GSD   | Ground Sampling Distance  |
| HDF-5 | Hierarchical Data Format Version 5  |
| IEC   | International Electrotechnical Commission   |
| IHO   | International Hydrographic Organization   |
| IMO   | International Maritime Organization   |
| ISO   | International Organization for Standardization  |
| PS    | Product Specification   |
| RMSE  | Root Mean Square Error  |
| SD    | Standard Deviation  |
| XML   | eXtensible Markup Language  |

## C-5 Overview of data quality measures

The relations between data quality terms are presented in Figure C-5-1 below. The terms have been defined in clause C-4.1.

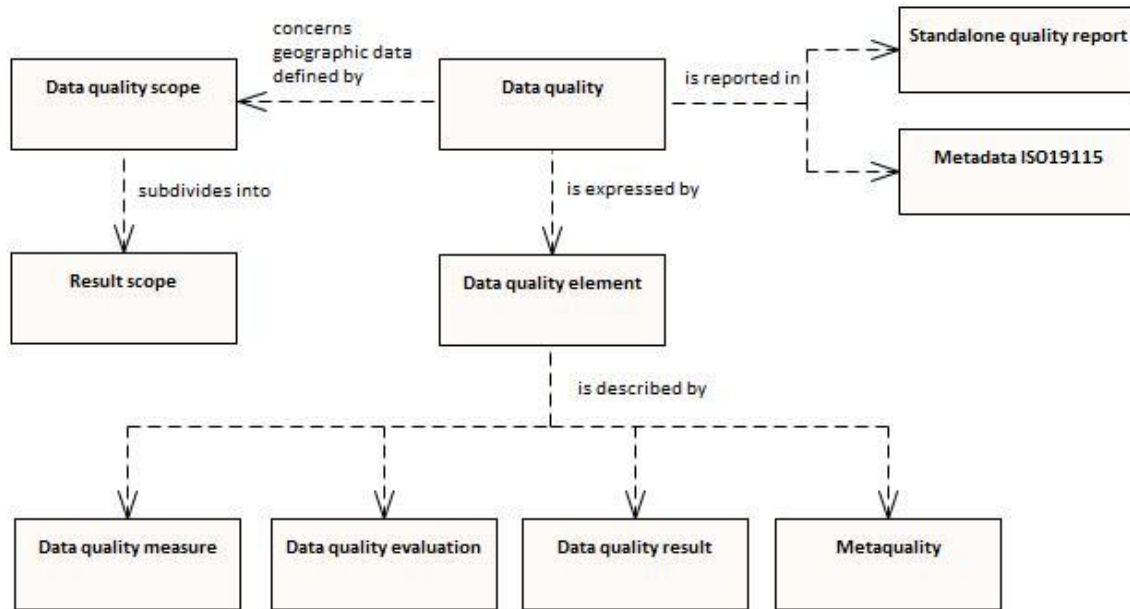


Figure C-5-4 – Conceptual model of quality for geographic data<sup>6</sup>

The components of Data Quality Measure can be divided into the following elements<sup>7</sup>:

1. Completeness
2. Logical Consistency
3. Positional Accuracy
4. Thematic Accuracy
5. Temporal Quality
6. Aggregation
7. Usability

Data Quality Evaluation can be divided into the following elements:

1. Data Quality Full Inspection
2. Data Quality Sample Based Inspection
3. Data Quality Indirect Evaluation
4. Data Quality Aggregation Derivation

Data Quality Result can be divided into the following elements:

1. Data Quality Conformance Result
2. Data Quality Quantitative Result
3. Data Quality Descriptive Result

<sup>6</sup> ISO 19157, Geographic Information – Data Quality, page 6.

<sup>7</sup> ISO 19157 – Geographic Information, Data Quality page 7.

## C-5.1 Data quality measures

The description of data quality measures in this section is quoted from ISO 19157.

**Completeness** is defined as the presence and absence of features, their attributes and relationships. It consists of two Data Quality Elements:

- Commission – excess data present in a dataset;
- Omission – data absent from a dataset.

**Logical Consistency** is defined as the degree of adherence to logical rules of data structure, attribution and relationships (data structure can be conceptual, logical or physical). If these logical rules are documented elsewhere (for example in a Product Specification) then the source should be referenced (for example in the data quality evaluation). It consists of four Data Quality Elements:

- Conceptual consistency – adherence to rules of the conceptual schema;
- Domain consistency – adherence of values to the value domains;
- Format consistency – degree to which data is stored in accordance with the physical structure of the dataset;
- Topological consistency – correctness of the explicitly encoded topological characteristics of a dataset.

**Positional Accuracy** is defined as the accuracy of the position of features within a spatial reference system. It consists of three Data Quality Elements:

- Absolute or external accuracy – closeness of reported coordinate values to values accepted as or being true;
- Relative or internal accuracy – closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true;
- Gridded data positional accuracy – closeness of gridded data spatial position values to values accepted as or being true.

**Thematic Accuracy** is defined as the accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships. It consists of three Data Quality Elements:

- Classification correctness – comparison of the classes assigned to features or their attributes to a Universe of Discourse (for example ground truth or reference data);
- Non-quantitative attribute correctness – measure of whether a non-quantitative attribute is correct or incorrect;
- Quantitative attribute accuracy – closeness of the value of a quantitative attribute to a value accepted as or known to be true.

**Temporal Quality** is defined as the quality of the temporal attributes and temporal relationships of features. It consists of three Data Quality Elements:

- Accuracy of a time measurement – closeness of reported time measurements to values accepted as or known to be true;
- Temporal consistency – correctness of the order of events;
- Temporal validity – validity of data with respect to time.

**Usability** is based on user requirements. All quality elements may be used to evaluate usability. Usability evaluation may be based on specific user requirements that cannot be described using the quality elements described above. In this case, the usability element must be used to describe specific quality information about a dataset's suitability for a particular application or conformance to a set of requirements.

## C-6 Recommendations for Product Specification developers

S-97 Part C applies the Data Quality concept from ISO 19157 to the development of S-100 based Product Specifications. It provides ten recommendations which assist in finding and applying applicable Data Quality Measures as described in S-100.

### C-6.1 Completeness > Commission / Omission

DQ\_CompletenessCommission: Defined in S-100 Appendix 4c-C, Hydrographic Quality Metadata Attribute Definitions.

DQ\_CompletenessOmission: Defined in S-100 Appendix 4c-C, Hydrographic Quality Metadata Attribute Definitions.

**Recommendation 1:** Data Quality Measure Completeness (Commission/Omission) to be included in the Product Specification.

### C-6.2 Logical Consistency > Conceptual Consistency

The Conceptual Schema Language is described in S-100 Part 1 – Conceptual Schema Language. It provides the description of:

- classes;
- attributes;
- basic data types;
- primitive types;
- complex types;
- predefined derived types;
- enumerated types;
- codelist types;
- relationships and associations;
- composition and aggregation;
- stereotypes and optional, conditional, and mandatory attributes and associations;
- naming and name spaces;
- notes; and
- packages.

**Recommendation 2:** Data Quality Measure Conceptual Consistency to follow the guidelines from S-100 Part 1 and to be included in the Product Specification.

### C-6.3 Logical Consistency > Domain Consistency

This is described in S-100 Part 5 – Feature Catalogue. This Part provides a standard framework for organizing and reporting the classification of real world phenomena in a set of geographic data. It defines the methodology for classification of the feature types and specifies how they are organized in a Feature Catalogue and presented to the users of a set of geographic data. This methodology is applicable to creating catalogues of feature types in previously uncatalogued domains and to revising existing Feature Catalogues to comply with standard practice. It applies to the cataloguing of feature types that are represented in digital form. Its principles can be extended to the cataloguing of other forms of geographic data.

**Recommendation 3:** Data Quality Measure Domain Consistency to follow the guidelines from S-100 Part 5 and to be included in the Product Specification.

### C-6.4 Logical Consistency > Format Consistency

This is described in S-100 Part 10 – Encoding formats. S-100 does not mandate particular encoding formats so it is left to developers of Product Specifications to decide on suitable encoding standards and to document their chosen format. The issue of encoding information is complicated by the range of encoding standards that are available, which include but are not limited to: ISO/IEC8211, GML, XML, GeoTiff, HDF-5, JPEG2000.

**Recommendation 4:** Data Quality Measure Format Consistency to follow the guidelines from S-100 Part 10 and to be included in the Product Specification.

### C-6.5 Logical Consistency > Topological Consistency

This is described in S-100 Part 7 – Spatial Schema. It supports 0, 1, 2, and 2.5 dimensional spatial schemas and two levels of complexity – geometric primitives and geometric complexes.

The conditions for topological consistency are provided in S-100 Part 7 – Spatial Schema, clause 7-4.3 and Appendix 7-A. The Figures in clause 7-4.3 and Appendix 7-A should be referred to for more details.

**Recommendation 5:** Data Quality Measure Topological Consistency to follow the guidelines from S-100 Part 7 and to be included in the Product Specification.

### C-6.6 Positional Accuracy

Positional Accuracy is described by S-100 Part 4c – Metadata - Data Quality.

This is further subdivided into Absolute or External Accuracy, Vertical Position Accuracy, Horizontal Positional Accuracy, Gridded Data Position Accuracy.

One should take notice of the different ways in which spatial data referencing is applied. Point set data includes a coordinate direct position for each point in the point set (points/curves). Gridded data references the grid as a whole. The two spatial properties of gridded data describe how the spatial extent was tessellated into small units and spatial referencing to the earth. The ISO 19123 Standard indicates that a grid may be defined in terms of a coordinate reference system. This requires additional information about the location of the grid's origin within the coordinate reference system; the orientation of the grid axes;

and a measure of the spacing between the grid lines. A grid defined in this way is called a rectified grid. If the coordinate reference system is related to the Earth by a datum, the grid is a georectified grid. The essential point is that the transformation of grid coordinates to coordinates of the external coordinate reference system is an affine transformation.

For Positional Accuracy, currently in the hydrographic community the 95% confidence level (Gaussian distribution) is commonly used. The Root Mean Square Error (RMSE) is commonly used in the scientific community. RMSE is the square root of the average of the set of squared differences between dataset coordinate values and coordinate values from an independent source of higher accuracy for identical points.

Other calculation methods are also possible, depending on the specific Product Specification. Comparisons of S-44, S-101, S-102, and S-121 were done. They use various methods of calculating and/or expressing the same concept (uncertainty). Some of the methods in use are:

- S-44: 95% (2\*SD)
- S-121: Standard Circular Error ( $=0.7071*\text{SQRT}(\text{SD}(X)+\text{SD}(Y))$ ), converted to a category attribute.
- S-101: All horizontal positional (2D), vertical (1D), horizontal distance (1D) and orientation (1D) uncertainty attributes concern the 95% confidence level of the variation associated with all sources of measurement, processing and visualization error. Uncertainty is encoded as: (1) uncertainty fixed (The maximum absolute value of the one-dimensional error (for vertical) or two-dimensional error (for horizontal). The error is assumed to be positive and negative.) and (2) uncertainty variable factor (The factor to be applied to a quantity to calculate its uncertainty. The fraction that equates to the factor (or percentage) contributing to the variable uncertainty component is indicated, that is a factor of 5% is encoded as 0.05.)

The situation above is like using nautical miles, imperial miles, and kilometers separately between different Product Specifications – which may work well in isolation, but when combining the different Product Specifications and use computer algorithms to create a “smart” system based on these quality parameters, great care must be taken in developing systems to ensure confusion is not introduced.

The IHO DQWG is considering different approaches to addressing this situation, and may leave the separate Product Specifications with their own parameters but inform developers and other users of the Specifications how to convert from one accuracy standard to another.

**Recommendation 6:** Data Quality Measure Positional Accuracy to follow the guidelines from S-100 Part 4c and to be included in the Product Specification. The calculation of the Positional Accuracy is to be further harmonized where possible with other S-100 based Product Specifications.

### C-6.7 Thematic Accuracy

Thematic Accuracy is described in ISO 19157.

This is further subdivided into: Quantitative Attribute Accuracy; Non-quantitative Attribute Accuracy; and Thematic Classification Correctness.

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. This document considers two types of object:

1. Features – features are defined together with their properties.

2. Information Types – information types are used to share information among features and other information types. Information types have only thematic attribute properties.

For Thematic Classification Correctness, the assignment of an item to a certain class can either be correct or incorrect.

For Quantitative Attribute Accuracy, the accuracy of quantitative attributes can be measured in terms of uncertainty intervals.

For Non-quantitative Attribute Accuracy, the attribute values of non-quantitative attributes can be correct or incorrect.

**Recommendation 7:** Data Quality Measure Thematic Accuracy to follow the guidelines from S-100 Part 4c and to be included in the Product Specification.

### C-6.8 Temporal Quality

Temporal Quality is described by Part 4c – Metadata - Data Quality.

Temporal Consistency and Temporal Validity are recommended to be included as this provides the user with the guarantee that any information in the temporal-spatial domain is registered correctly. For data elements with a very precise temporal attribute (for example remote sensing), the temporal accuracy may also be provided.

**Recommendation 8:** Data Quality Measure Temporal Quality to follow the guidelines from S-100 Part 4c and to be included in the Product Specification. Temporal Consistency and Temporal Validity should be included.

### C-6.9 Aggregation

Data Quality specified at upper level (for example series) is applicable at lower level (for example dataset), see Table C-6-1 (quoted from ISO 19157) below. If the Data Quality differs between the upper and lower level, then supplemental information should be provided at the lower level.

|  |                  |                    |
|--|------------------|--------------------|
| <p><b>Upper level</b></p> <p>↓</p> <p><b>Lower level</b></p> | Series           |                    |
|  | Dataset          |                    |
|  | Subset           |                    |
|  | Feature type     | Attribute type     |
|  | Feature instance | Attribute instance |

*Table C-6-5 - Hierarchical levels*

In the case of aggregation of different quality results, the standalone quality report should provide full information on the original results (with evaluation procedures and measures applied); the aggregated result; and the aggregation method, whereas the metadata may describe only the aggregated result with a reference to the original results described in the standalone quality report.

The aggregated Data Quality result provides a result if the dataset has passed conformance to the Data Product Specification.

**Recommendation 9:** Data Quality Measure Aggregation results should be included to indicate if the dataset/dataset series have passed the Product Specifications.

### C-6.10 Usability

Usability is based on user requirements. All quality elements may be used to evaluate usability. Usability evaluation may be based on specific user requirements that cannot be described using the quality elements described above. In this case, the usability element shall be used to describe specific quality information about a dataset's suitability for a particular application or conformance to a set of requirements.

All Product Specifications should have a paragraph describing Data Quality. To ensure harmonization across different Product Specifications, DQWG recommends that all Product Specifications share a common text explaining the concept of Data Quality -> Introduction to Data Quality. The text below is a proposal for this common introduction:

#### Introduction to data quality

Data quality allows users and user systems to assess fitness for use of the provided data. Data quality measures and the associated evaluation are reported as metadata of a data product. This metadata improves interoperability with other data products and provides usage by user groups that the data product was not originally intended for. The secondary users can make assessments of the data product usefulness in their application based on the reported data quality measures.

For <*this Product Specification*> the following Data Quality Elements have been included<sup>8</sup>:

- Conformance to this Product Specification;
- Intended purpose of the data product;
- Completeness of the data product in terms of coverage;
- Logical Consistency;
- Positional Uncertainty and Accuracy;
- Thematic Accuracy;
- Temporal Quality;
- Aggregation measures;
- Validation checks or conformance checks including:
  - o General tests for dataset integrity;
  - o Specific tests for a specific data model.

**Recommendation 10:** Paragraph "Introduction to data quality" to be used as a template by all S-100 based Product Specifications.

---

<sup>8</sup> As deemed necessary by the IHO – Hydrographic Standards and Services Committee



## C-7 Data quality measures

This list is derived from ISO 19157.

There are twenty different data quality measures defined that can be used for validation of S-100 based Product Specifications. These are shown in Table C-7-1 below. There is a recommendation following the table for the target result of Positional Accuracy for depth contour lines and gridded bathymetry.

The full list of ISO 19157 measures applicable to S-100-based Product Specifications is available in S-100 Appendix 4c-C – Hydrographic Quality Metadata Attribute Definitions.

| Data Quality Measure                          | Definition  | DQ measure / description  | Evaluation scope                           | Applicable to spatial representation types |
|---|---|---|--|--|
| Completeness / Commission                     | Excess data present in a dataset, as described by the scope.  | numberOfExcessItems / This data quality measure indicates the number of items in the dataset, that should not have been present in the dataset.   | dataset/dataset series                     | All S-100 based PS                         |
| Completeness / Commission                     | Excess data present in a dataset, as described by the scope.  | numberOfDuplicateFeatureInstances / This data quality measure indicates the total number of exact duplications of feature instances within the data.  | dataset/dataset series                     | All S-100 based PS                         |
| Completeness / Omission                       | Data absent from the dataset, as described by the scope.  | numberOfMissingItems / This data quality measure is an indicator that shows that a specific item is missing in the data.  | dataset/dataset series/spatial object type | All S-100 based PS                         |
| Logical Consistency / Conceptual Consistency  | Adherence to the rules of a conceptual schema.  | numberOfInvalidSurfaceOverlaps / This data quality measure is a count of the total number of erroneous overlaps within the data. Which surfaces may overlap and which must not is application dependent. Not all overlapping surfaces are necessarily erroneous.  | spatial object / spatial object type       | PS with geometric surfaces.                |
| Logical Consistency / Domain Consistency      | Adherence of the values to the value domains.   | numberOfNonconformantItems / This data quality measure is a count of all items in the dataset that are not in conformance with their value domain.  | spatial object / spatial object type       | All S-100 based PS                         |
| Logical Consistency / Format Consistency      | Degree to which data is stored in accordance with the physical structure of the data set, as described by the scope | physicalStructureConflictsNumber / This data quality measure is a count of all items in the dataset that are stored in conflict with the physical structure of the dataset.   | dataset/dataset series                     | All S-100 based PS                         |
| Logical Consistency / Topological Consistency | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.        | rateOfFaultyPointCurveConnections / This data quality measure indicates the number of faulty link-node connections in relation to the number of supposed link-node connections. This data quality measure gives the erroneous point-curve connections in relation to the total number of point-curve connections. | spatial object / spatial object type       | PS with curves.                            |
| Logical Consistency / Topological Consistency | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.        | numberOfMissingConnectionsUndershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to undershoots.   | spatial object / spatial object type       | PS with curves.                            |

| Data Quality Measure                                 | Definition   | DQ measure / description  | Evaluation scope                           | Applicable to spatial representation types                             |
|--|--|---|--|--|
| Logical Consistency / Topological Consistency        | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope. | numberOfMissingConnectionsOvershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots.   | spatial object / spatial object type       | PS with curves.  |
| Logical Consistency / Topological Consistency        | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope. | numberOfInvalidSlivers / This data quality measure is a count of all items in the dataset that are invalid sliver surfaces. A sliver is an unintended area that occurs when adjacent surfaces are not digitized properly. The borders of the adjacent surfaces may unintentionally gap or overlap to cause a topological error. | dataset / dataset series                   | PS with geometric surfaces.  |
| Logical Consistency / Topological Consistency        | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope. | numberOfInvalidSelfIntersects / This data quality measure is a count of all items in the dataset that illegally intersect with themselves.  | spatial object / spatial object type       | PS with curves / geometric surfaces.                                   |
| Logical Consistency / Topological Consistency        | Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope. | numberOfInvalidSelfOverlap / This data quality measure is a count of all items in the dataset that illegally self-overlap.  | spatial object / spatial object type       | PS with curves / geometric surfaces.                                   |
| Positional Accuracy / Absolute or External Accuracy  | Closeness of reported coordinative values to values accepted as or being true.                               | Root Mean Square Error / Standard deviation, where the true value is not estimated from the observations but known a priori.  | spatial object / spatial object type       | PS with objects that have coordinative values associated.              |
| Positional Accuracy / Vertical Position Accuracy     | Closeness of reported coordinative values to values accepted as or being true.                               | linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.  | spatial object / spatial object type       | PS with objects that have a vertical coordinative values associated.   |
| Positional Accuracy / Horizontal Position Accuracy   | Closeness of reported coordinative values to values accepted as or being true.                               | linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.  | spatial object / spatial object type       | PS with objects that have a horizontal coordinative values associated. |
| Positional Accuracy / Gridded Data Position Accuracy | Closeness of reported coordinative values to values accepted as or being true.                               | Root mean square error of planimetry / Radius of a circle around the given point, in which the true value lies with probability P.  | spatial object / spatial object type       | PS with objects that have a gridded coordinative values associated.    |
| Temporal Quality / Temporal Consistency              | Consistency with time.   | Correctness of ordered events or sequences, if reported.  | dataset/dataset series/spatial object type | PS with objects that have a time value associated.                     |

| Data Quality Measure                                   | Definition  | DQ measure / description  | Evaluation scope                           | Applicable to spatial representation types   |
|--|---|---|--|--|
| Thematic Accuracy / ThematicClassification Correctness | Comparison of the classes assigned to features or their attributes to a universe of discourse.                  | miscalculationRate / This data quality measure indicates the number of incorrectly classified features in relation to the number of features that are supposed to be there. [Adapted from ISO 19157] This is a RATE which is a ratio, and is expressed as a REAL number representing the rational fraction corresponding to the numerator and denominator of the ratio. For example, if there are 1 items that are classified incorrectly and there are 100 of the items in the dataset then the ratio is 1/100 and the reported rate = 0.01. | dataset/dataset series/spatial object type | All S-100 based PS.  |
| Aggregation Measures / AggregationMeasures             | In a data Product Specification, several requirements are set up for a product to conform to the Specification. | DataProductSpecificationPassed / This data quality measure is a boolean indicating that all requirements in the referred data Product Specification are fulfilled.  | dataset/dataset series/spatial object type | PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types |
| Aggregation Measures / AggregationMeasures             | In a data Product Specification, several requirements are set up for a product to conform to the specification. | DataProductSpecificationFailRate / This data quality measure is a number indicating the number of data Product Specification requirements that are not fulfilled by the current product/dataset in relation to the total number of data Product Specification requirements.   | dataset/dataset series/spatial object type | PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types |

Table C-7-1 – Recommended data quality measures

NOTE: Recommendations for Positional Accuracy / Absolute or External Accuracy<sup>9</sup>:

Maximum RMSE (horizontal) =  $E / 10000$

Maximum RMSE (vertical) =  $Vint / 6$

Recommendation for Positional Accuracy / Gridded Data Position Accuracy:

Maximum RMSE (horizontal) =  $GSD / 6$

Maximum RMSE (vertical) =  $GSD / 3$

Where:

E = Denominator of intended scale of mapping

Vint = normal contour line interval

GSD = Ground Sampling Distance

## C-8 Minimum standard for data validation

A minimum standard set of checks for data validation is under development and will be added to this document when completed.

<sup>9</sup> INSPIRE D2.8.II.1 Data Specification on Elevation – Technical Guidelines, page 95

Page intentionally left blank