IMO EGDH comparison of IMO Compendium data reference model and IHO S-100

Submitted by:	NIPWG, Vice Chair (IHO EGDH rep)
Executive Summary:	Comparison of IMO Compendium and S-100
Related Documents:	-
Related Projects:	IMO EGDH collaboration, S-131

Introduction / Background

IMO Expert Group on Data Harmonization (EGDH) is a subgroup of the IMO FAL committee. The EGDH maintains the IMO Compendium on Facilitation and Electronic Business. Expansion of the IMO Compendium beyond the FAL convention has introduced some overlap with IHO GI-Registry and e-Navigation concepts.

An overview of EGDH collaboration was given at HSSC 15 and NIPWG VTC last summer; https://iho.int/uploads/user/Services%20and%20Standards/NIPWG/NIPWG%20VTC%202023/NIPWG_VTC02_2023_03.1A_EN_EGDH.pdf

Since all e-navigation related Maritime Services should be conformant with the S-100 framework, while some of the Maritime Services relevant to the FAL Committee use the IMO reference data model, also identifying some structural differences might be necessary to support the harmonization of specific data elements. EGDH 8 agreed to initiate a mapping exercise of selected IMO datasets into S-100 structure, in order to identify (structural) differences. The mapping exercise was conducted after EGDH 8 by the IHO NIPWG representative. Progress and findings of the exercise will be presented at NIPWG 10, 09/2023 and EGDH 9 10/2023.

A mapping exercise was conducted during the summer 2023, where the IMO Compendium reference model and datasets were transferred into a S-100 style Feature Catalogue, a test data set generated and structural findings reported. Only data publicly available at https://imocompendium.imo.org/ was used.

Attached to this input is a zip-file containing the;

- Mapping exercise draft report (PDF)
- Generated Feature Catalogue XML (S-100 4.0.0 FC schema) (XML and PDF)
- Test Dataset FAL form 6 Passenger list printed as a GML 5.0 dataset (GML and PDF)

Additional to being a practical learning process, the mapping exercise aimed to verify whether the IMO Compendium data could fit a S-100 style structure. Further, in order to facilitate harmonizing and mapping of data elements between the IMO Compendium and IHO registry / IHO S-100 Product specifications it was assumed documentation of structural differences might be beneficial.

The mapping exercise could also provide input for some discussions at EGDH 8 and following intersessional work, regarding the IMO Compendium development;

- Possible need for a mechanism to enforce optional / mandatory attributes
- Possible need for sub-models (complex structures)
- Possible need for additional user guidance
- Identifiers for berth, terminal, port etc. (IMO JIT-dataset / IHO S-131)

Analysis/Discussion

Machine readability

The EXCEL-version of IMO Compendium was converted into CSV- datafiles, then imported into an S-100 structure using quite low-level custom scripting. Enhanced machine-readability of the publicly available data might aid developers in adopting and updating implementations based on the compendium.

S-100 unique element identifiers

S-100 uses "camelCase" of element names as unique identifiers. IMO Compendium use numbers for each Data element (IMO 0001, IMO 0002 etc.) as unique identifier. Names of data elements in the IMO compendium had to be parsed in order to fit S-100. Sometimes S-100 naming requirements demanded renaming duplicate names from IMO reference model.

Enumerations (IMO Codelists)

Enumeration items has to be added to an S-100 Feature Catalogue. The IMO Compendium does not contain all enumerations (codelists) within the Compendium, and these has to be manually added for an S-100 implementation.

IMO Datasets are Data Dictionaries

IMO Datasets are dictionaries of data elements. The Datasets does not describe structure, multiplicity or indication of mandatory or optional attributes. The S-100 implementation of datasets had to add at least the minimum complexity needed to support multiplicities.

The example used in the mapping exercise was the Passenger list, where one list contains multiple passengers. Parts of the complex structure defined in the IMO reference model (UML) had to additionally be cross-referenced and implemented.

With a background in S-100 data modelling and object oriented programming, the need to cross-reference the Datasets (Dictionaries of Data Elements) to the separate Reference model, where the objects are grouped into a top-down hierarchy, posed a bit of a challenge.

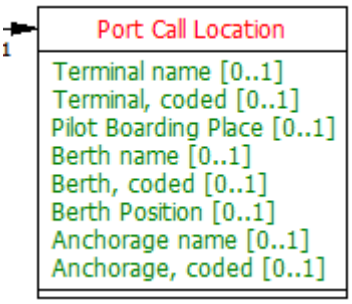
Attribute encapsulation

The main structural difference found between the models, seem to be that IMO Datasets only reference single Data Elements. (In S-100 the equivalent would be to only support a Simple Attribute.) There is no equivalent of S-100 Complex attributes (=EDIFACT Composites) in the IMO Compendium. Despite the grouping of attributes in the reference model, the composites (groups / classes / objects) (example Port Call Location in the figure on next page) cannot be directly referenced as a group in a Dataset. Only individual attributes (Data elements) are referenced.

Allowing reference to composite Data elements makes attribute encapsulation possible. When (encapsulated) attributes are accessed only through the container and not directly, the model itself provides context to the attribute. For example. a generic attribute "Familyname" within a "Passenger" object could be assumed to provide the name of that Passenger.

Conclusions

The use of complex structures and encapsulated attributes in S-100 but not in IMO Compendium affects the mapping of elements between the IMO Compendium and IHO, especially if referencing the IHO registry only. This should be noted such that it is clear a reference to a given Product specification is needed, in order to ensure usage of a given concept within a Product specification.



For example, S-131 does support encoding of names for Terminal and Berth, but there are no separate attributes with a name similar to *Terminal name*, *Berth name* or *Anchorge name.* In S-131 *Terminal, Berth and Anchorage* are FeatureTypes. Each FeatureType contain an encapsulated attribute FeatureName, to describe the name of that current feature.

Similarly, IHO registry contain attributes for encoding identifiers such as FacilityNumber or UNLOCODE, but only S-131 Product specification reveals, whether these attributes are available, optional or mandatory in a certain FeatureType such as Terminal or Berth.

The addition of Complex structure / Composites to IMO Datasets might solve some of the development needs identified by EGDH, such as a need for sub-models, managing optional / mandatory attributes and map identifiers for berth, terminal etc.

IHO S-100 is a complex and multi- layered structure. All concepts are defined in the registry, but the use of these concepts and especially the encapsulated attributes available could differ between Product specifications. Mapping of some data elements against IHO might therefore need to refer to encapsulated attributes within a complex structure in a given Product specification, instead of a direct reference to IHO registry.

Identifiers for IHO elements should refer the "camelCase" encoding of the name, which is a unique identifier across all levels. Sometimes also the containing object must be referred to provide context.

Recommendations
To be discussed

Justification and Impacts
-

Action Required of NIPWG
The NIPWG is invited to:

a. note this paper
b discuss and consider recommendations