IMO Compendium S-100 mapping exercise

Draft report for NIPWG 10
SE 4.9.2023

Contents

1 IMO EGDH

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.

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.

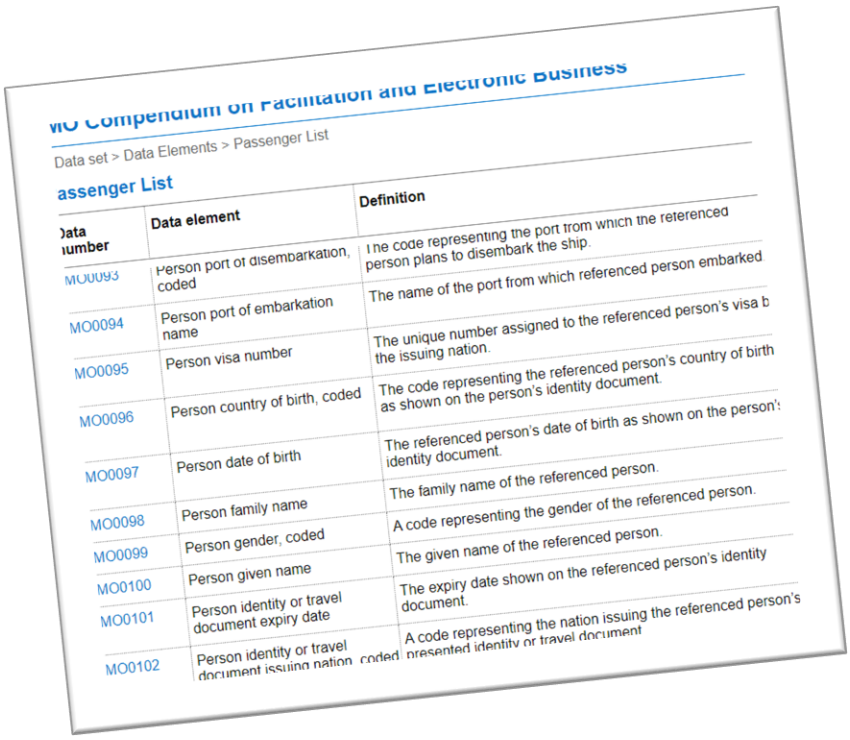
The following discussion points, related to this IMO Compendium and this mapping exercise have been identified during EGDH 8 and following intersessional work ;

- Need for a mechanism to enforce optional / mandatory attributes
- Possible need for sub-models (complex structures)
- Need for additional user guidance
- Use of identifiers for berth, terminal, port etc. (JIT dataset)

2 IMO Compendium

The IMO Compendium consists of an IMO Reference Data Model, and several IMO Data Sets. The building blocks of Data sets are Data Elements. The IMO Compendium supports electronic data exchange of information required for port calls. It is recommended for use by Maritime Single Window- systems. The initial Data sets included electronic versions of the FAL- convention standard forms for reporting. Currently, the Compendium is expanding to areas beyond the FAL Convention, including exchange of logistics and operational port and shipping data.

The IMO Compendium follows UN/EDIFACT (the United Nations rules for Electronic Data Interchange for Administration, Commerce and Transport), syntax and procedures. Most organizations involved in EGDH-work are using UN/EDIFACT for data transmission. UN/EDIFACT is escribed as a set of internationally agreed standards, directories, and guidelines for the electronic interchange of structured data, between independent computerized information systems.



IMO Compendium on Facilitation and Electronic Business

Data set > Data Elements > Passenger List

assenger List

Data number	Data element	Definition
MO0093	Person port of disembarkation, coded	The code representing the port from which the referenced person plans to disembark the ship.
MO0094	Person port of embarkation name	The name of the port from which referenced person embarked
MO0095	Person visa number	The unique number assigned to the referenced person's visa b the issuing nation.
MO0096	Person country of birth, coded	The code representing the referenced person's country of birth as shown on the person's identity document.
MO0097	Person date of birth	The referenced person's date of birth as shown on the person's identity document.
MO0098	Person family name	The family name of the referenced person.
MO0099	Person gender, coded	A code representing the gender of the referenced person.
MO0100	Person given name	The given name of the referenced person.
MO0101	Person identity or travel document expiry date	The expiry date shown on the referenced person's identity document.
MO0102	Person identity or travel document issuing nation, coded	A code representing the nation issuing the referenced person's presented identity or travel document

Figure 1 In 2023, the Compendium contains around 500 data elements, used in 17 datasets.

The compendium should not be considered a convention or regulation by itself, but merely a tool to support data exchange. Member states are encouraged by FAL 46 to use the IMO Compendium as a reference for implementations of the Maritime Single Window. The IMO Compendium is available online as an HTML or an EXCEL- spreadsheet at https://imocompendium.imo.org/. The compendium is described as a reference manual for harmonizing systems that handle declarations required by the FAL Convention.

2.1.1 ISO 280005

ISO technical committee ISO/TC 8, Ships and marine technology, has been working with Electronic Port Clearance. This work is related to the IMO Compendium. ISO 28005:2 dictionary of Core data elements include a mapping of IMO Compendium Data elements.

2.1.2 Electronic port clearance (EPC) – Part 1: Message structures

ISO 28005-1:20213 defines XML message structures for the transmission of information between a ship or its representatives and certain organizations responsible for the processing of the ship's port clearance request. The information intended to be transferred is that which is defined by the FAL Convention and other related international instruments as identified by ISO 28005-2. These message structures are primary intended for machine -to- machine data transfers.

2.1.3 Electronic port clearance (EPC) – Part 2: Core data elements

ISO 28005-2:2021 provides the definition of core data elements for use in electronic port clearance (EPC) messages. It does not define any structuring of messages nor provides any guidance on what information is required for a particular purpose; it is a general data dictionary for safety, security or operation-related maritime information. The specifications in this document is compatible with the definitions in the IMO Reference Data Model and the mapping between ISO 28005 and the data element list in the IMO Reference Data Model is defined in Annex B.

2.1.4 Electronic port clearance (EPC) — Part 3: Technical standard for data exchange

ISO 28005-3 (under development) defines basic message exchanges and data elements used in the coordination of actors at sea and land in conjunction with a ship's port call. The scope covers vessel notifications and declarations, movements, geographical positions, and time stamps. The standard will define "skeleton" processes to provide the context for the message exchanges. An important part of this is to provide information and proposals to the port and shipping sectors as well as port and maritime administrations on how to facilitate JIT (Just in Time) arrival of ships.

3 S-100 overlap between JIT- data and IHO S-131

A 2020 partnership agreement between ISO, IMO, the World Customs Organization (WCO) and UNECE aims to expand the IMO Reference Data Model beyond the Facilitation Convention, such as exchanging operational data that could help facilitate just-in-time operation of ships. Just-in-time operation allows ships to optimize their speed, so they arrive at their destination port when their berth is ready for them, thereby saving energy and cutting costs and emissions.

Figure 2Figure 2 Image presented during IMO Maritime Single Window Symposium 18-19 January 2023. Ship

Overlap of IMO / IHO data was initially identified to occur at ports, where operational and administrative data is related to nautical data presented on charts. The main overlap is found between the JIT-dataset and IHO S-131, where JIT data refer to Terminals, Berths and other specific features described in S-131.

4 Mapping exercise

The main aim of the exercise was to produce information to support mapping and harmonization of individual Data elements. By doing a mapping of some selected data, the structural differences needed to consider should become apparent. Another aim was to investigate, whether IMO Compendium Datasets could be implemented using S-100 structures (Feature Catalogue). The aim of the exercise was not to initiate the creation of a new S-100 Product specification.

First, a comparison of the models was conducted. The IMO compendium is a reference model, containing guidance for implementation. S-100 consist of several parts, with reference structure, registries and data models of product specifications. The machine-readable and formal description of actual S-100 product specifications as Feature Catalogues could be considered an implementation. ISO 28005 standards are not included in this comparison, even though these seem to cover some of the functionality included in S-100.

4.1 Step 0 - Comparison of models

IMO Compendium	IHO S-100	Notes
Reference data model The IMO Reference data model describes a common hierarchical structure of all the available data elements. EDIFACT has a hierarchical structure where the top level is referred to as an *interchange*, and lower levels contain multiple *messages* which consist of *segments*, which in turn consist of *composites*. The final iteration is an *element*.	**S-100 framework / Registry** S-100 defines the framework for creation of S-100 products. Each Product specification additionally defines the data model of the product. Product specifications re-use elements from the IHO registry, and usually structures from existing specifications, but the hierarchical structure is individual for each product specification.	EDIFACT composites are not used in the IMO Compendium
Data element The IMO Compendium consist of Data elements. Each Data element has a name, definition and a unique Data number taking the form of 'IMO0001', ' IMO0002' etc.	**S-100 Concept** In S-100, each object (concept) has a unique name (camelCase) and a definition, but no separate data number.	
Data element Although the IMO reference model (UML) shows containers of attributes, these containers are not enforced within the datasets. IMO Datasets does not contain "composites" or another S-100 equivalent to complex attributes.	**Simple / Complex attributes** S-100 defines attributes as either simple or complex. A simple attribute carries a specific value such as a number, string or date. A complex attribute is an aggregation of other attributes, either simple or complex. A complex attribute is thus a container of other attributes.	IMO Compendium does not contain composites / complex attributes.

Data format The allowed format is described using an alphanumeric code syntax adopted from UN/EDIFACT. The format descriptor can define alphanumeric or numeric characters, or a combination of both. Additionally, either fixed or maximum number of characters and precision for decimal values can be defined. AS an example, for dates, the format descriptor 'an..35' is used. This would allow for any arbitrary string of maximum length 35 characters.	**Primitive types** S-100 main document defines the primitive data types that are supported in S-100. These include *integer, real, boolean, text, date, time, datetime* and a few special cases of these.	
Codelist Data elements of type Codelist are restricted to only accept the specific values found in a list. Codelists include the identification of the DataElements to be restricted, and are defined as a table where each entry has a code, name and definition. The code (key) of the enumeration can take several forms, including alphanumeric characters, commas and special characters. External codelists are referenced, but not included in the Compendium itself.	**Enumeration / codelist** S-100 support similar use of enumerations / codelists, although the S-100 key (code) is always an integer value. S-100 demands that all enumerations / codelists and allowable enumeration items are defined within the product specification itself, or machine readable through an external service.	S-100 implementation demands all elements including enumeration items to be included in the Feature Catalogue.

Datasets IMO Datasets are specified subsets of the data within the IMO Reference Model. The named Datasets are mainly similar in structure to the declarations required by the FAL Convention, with a few additions. As a note, the classes (containers of data elements), as described in the hierarchical view of the IMO Reference data model, are not reproduced in the Dataset descriptions. Instead, only the actual Data elements (Simple attributes) are listed within each Dataset, in a flat-file format.	**FeatureTypes** S-100 Feature Catalogues are used to describe the data model of each Product specification. FeatureTypes consist of simple- and complex attributes and associations to other Feature- or InformationTypes. The equivalent of an IMO Dataset is considered a FeatureType, associated to other features.	IMO Datasets does not contain hierarchy, these are flat-file "dictionaries" of data elements. The containers described in the IMO reference model are not objects with encapsulated attributes.
Business rules Business Rules are recommended procedures for satisfying specific data element formats and requirements. Business rules are mapped to specific Data Elements.	**DCEG / Feature catalogue** S-100 Product specifications contain a separate DCEG (Data Capture and Encoding Guide), containing recommendations related to the data encoding. Data format restrictions are primarily enforced by restrictions within the data model.	

4.2 Step 1 – Importing IMO Compendium data

As S-100 Feature Catalogues are used to describe the data model of each Product specification, an attempt was made to generate an S-100 Feature Catalogue based on the IMO- compendium. The IMO Compendium FAL 46 EXCEL file version October 24 2022 was used. From the EXCEL- file, the worksheets including the *Structure, Codelists* and *BusinessRules* were used. These pages were deemed to contain the needed data for testing. The IMO Datasets were translated into S-100 FeatureTypes and a single FeatureCatalogue was generated for the IMO compendium.

For tis exercise, only the data publicly available online was used. Quite low-level scripting was needed to read selected data exported into CSV-data and create a single S-100 FeatureType for each IMO Datase.,The associated Complex- and Simple attributes were created by cross referencing the *reference model* and d*ata elements*. Available codelist-values were applied to Simple attributes and business rules added to attribute definitions.

The result was a generated S-100 Product specification Feature Catalogue XML- file. Slight modification was made, and the S100FC.XML validated against S-100 4.0.0. schema.

4.2.1 Data elements implemented as S-100 Simple attributes

Data elements are implemented as S-100 Simple attributes. Simple attributes are the basic building blocks of a S-100 data model. The IMO compendium format encoding was translated into primary types string, real or integer. Business rules are added as text to the attribute definition. The descriptive name of the attribute is used as the unique identifier, and the IMO- code added for reference to the definition. Some tweaking of the names was needed, and contrary to the S-100 convention, camel-case naming is not used for the imported data. Instead, the imported data uses underscore to separate words. CamelCase format can then be used later in order to distinguish between imported and amended structures.

Example of a name using underscore; dangerous_goods_carried_indicator

Example of S-100 camel-case naming of the same element; dangerousGoodsCarriedIndicator

4.2.2 Codelists

Most IMO Compendium codelists are referenced to external sources. In these cases the items are initially not imported. The codelists contained internally within the IMO Compendium are imported as S-100 enumeration values. In S-100, an enumeration is a type of Simple attribute, where the allowed values are restricted to those in the codelist.

```
<!--
**********************************************
3/419 S100_FC_SimpleAttribute: IMO0002 (text)  used in: agent_at_port_c, General_Declaration, Cargo_Declaration, Ships_Stores_Declaration,
Crews_Effects_Declaration, Crew_List, Passenger_List, Dangerous_Goods_Manifest, Security_Report, Ship_Reporting_Systems
**********************************************
-->
<S100FC:S100_FC_SimpleAttribute>
  <S100FC:name>IMO0002</S100FC:name>
  <S100FC:definition>The agent's employee contact name. The format to be used is: FamilyName, GivenName</S100FC:definition>
  <S100FC:code>agent_contact_name</S100FC:code>
  <S100FC:valueType>text</S100FC:valueType>
</S100FC:S100_FC_SimpleAttribute>
<!--
**********************************************
4/419 S100_FC_SimpleAttribute: IMO0003 (text)  used in: address_c, General_Declaration, Dangerous_Goods_Manifest, Security_Report, Stowaways,
Ship_Reporting_Systems
**********************************************
-->
<S100FC:S100_FC_SimpleAttribute>
  <S100FC:name>IMO0003</S100FC:name>
  <S100FC:definition>A code representing the country address of the referenced agent at port. ISO 3166-1 alpha-2 code (UN/EDIFACT Codes 3207)</
S100FC:definition>
  <S100FC:code>agent_country_code</S100FC:code>
  <S100FC:valueType>text</S100FC:valueType>
</S100FC:S100_FC_SimpleAttribute>
```

Figure 3 Simple attributes (Data elements IMO0002 and IMO0003) imported into an S-100 style FeatureCatalogue XML-file. The external codelist is not implemented, as this data is not available in the Compendium. This would have to be added separately or referenced to an existing ISO-3166-1 codelist.

4.2.3 IMO reference model implemented as a structure of Complex attributes

The IMO reference model is a top-down hierarchy, where Data elements are grouped into classes. Additionally to the separate UML- image of the reference model, also the path of the individual data-elements within the hierarchy is available in the EXCEL- workbook. This information is used to create a recursive structure of S-100 Complex attributes, that implements the hierarchy expressed in the UML. Complex attributes are given an additional postfix '_c' to the name, to identify the Complex attributes forming the IMO Reference model.

In S-100, all positions are geometries of features. Features can be associated to other features. Therefore, complex attributes containing sub-attributes such as latitude or longitude, has later to be converted into associated features.

```
*********************************************
13/72 S100_FC_ComplexAttribute: contact_details_c (contact_details_c)  used in: maritime_service_c
*********************************************
-->
<S100FC:S100_FC_ComplexAttribute>
  <S100FC:name>contact_details_c</S100FC:name>
  <S100FC:definition>contact_details_c</S100FC:definition>
  <S100FC:code>contact_details_c</S100FC:code>
  <S100FC:subAttributeBinding sequential="false">
    <S100FC:multiplicity>
      <S100Base:lower>0</S100Base:lower>
      <S100Base:upper xsi:nil="false" infinite="false">1</S100Base:upper>
    </S100FC:multiplicity>
    <S100FC:attribute ref="service_provider_contact_name"/>
  </S100FC:subAttributeBinding>
  <S100FC:subAttributeBinding sequential="false">
    <S100FC:multiplicity>
      <S100Base:lower>0</S100Base:lower>
      <S100Base:upper xsi:nil="false" infinite="false">1</S100Base:upper>
    </S100FC:multiplicity>
    <S100FC:attribute ref="communication_c"/>
  </S100FC:subAttributeBinding>
</S100FC:S100_FC_ComplexAttribute>
```

Figure 4 A complex attribute defined in the S-100 FeatureCatalogue. This attribute contains references to Simple attribute service_provider_contact_name and Complex attribute communication_c.

4.2.4 IMO Datasets implemented as Features

The Datasets (declarations / reports) are initially imported as S-100 Features. This initial choice was made in order to keep the structure as simple as possible at this stage. Normally features contain geographical positions, but a feature can also exist without a position.

The following Features were added from the FAL 46 version of the compendium;

"General Declaration", "Cargo Declaration", "Ships Stores Declaration", "Crews Effects Declaration", "Crew List", "Passenger List", "Dangerous Goods Manifest", "Security Report", "Waste Delivery in Port", "Maritime Declaration Health", "Just In Time Concept", "Stowaways", "Ship and Company Certificates", "Acknowledgement Receipt", "Maritime Service", "Ship Registry & Comp. Details", "Inspections", "PSC Inspection History Data", "Ship Reporting Systems"

The Datasets are initially implemented similarly to the structure described in the Compendium. This means, that the individual data elements are directly referenced in each Dataset with a default multiplicity indication of [0..1]. This means that each attribute is optional and maximum one instance is allowed. Especially mandatory attributes, and attributes where multiple instances are allowed must be further considered.

Cargo_Declaration

agent_contact_name	(IMO0002) text	[0..1]
agent_identification_number	(IMO0007) text	[0..1]
agent_name	(IMO0010) text	[0..1]
arrivaldeparture_code	(IMO0013) text	[0..1]
authentication_date	(IMO0014) text	[0..1]
authenticator_name	(IMO0016) text	[0..1]
authenticator_party_identification_number	(IMO0017) text	[0..1]
transport_equipment_identification_number	(IMO0021) text	[0..1]
cargo_item_description_of_goods	(IMO0022) text	[0..1]
cargo_item_gross_volume	(IMO0023) real	[0..1]
cargo_item_gross_weight	(IMO0024) real	[0..1]
cargo_item_hs_code	(IMO0025) text	[0..1]
cargo_item_marks_and_numbers	(IMO0026) text	[0..1]
cargo_item_number_of_packages	(IMO0028) integer	[0..1]
cargo_item_package_type_coded	(IMO0029) text	[0..1]
port_of_discharge_coded	(IMO0113) text	[0..1]
port_of_discharge_name	(IMO0114) text	[0..1]
port_of_loading_coded	(IMO0117) text	[0..1]
port_of_loading_name	(IMO0118) text	[0..1]
authenticator_role_coded	(IMO0128) text	[0..1]
ship_flag_state_coded	(IMO0138) text	[0..1]
ship_imo_number	(IMO0140) text	[0..1]
ship_name	(IMO0142) text	[0..1]
ship_stay_reference_number	(IMO0153) text	[0..1]
transport_contract_number	(IMO0170) text	[0..1]
voyage_number	(IMO0191) text	[0..1]
message_function_code	(IMO0305) text	[0..1]

Figure 5 Example of a Dataset as imported. At least indication of mandatory content and identifying where multiple instances of attributes are allowed has to be further considered.

4.2.5 Notes of the import exercise

The import resulted in a FeatureCatalogue- file containing Simple attributes, Complex attributes and FeatureTypes. As the compendium separates the reference model from the datasets, the Complex attributes reflecting the Reference model are not used within the datasets at this point.

Multiple *Communication*- classes existed in the reference model. These had to be uniquely renamed by prefixing. (service contact, dangerous goods, service, security officer, agent)

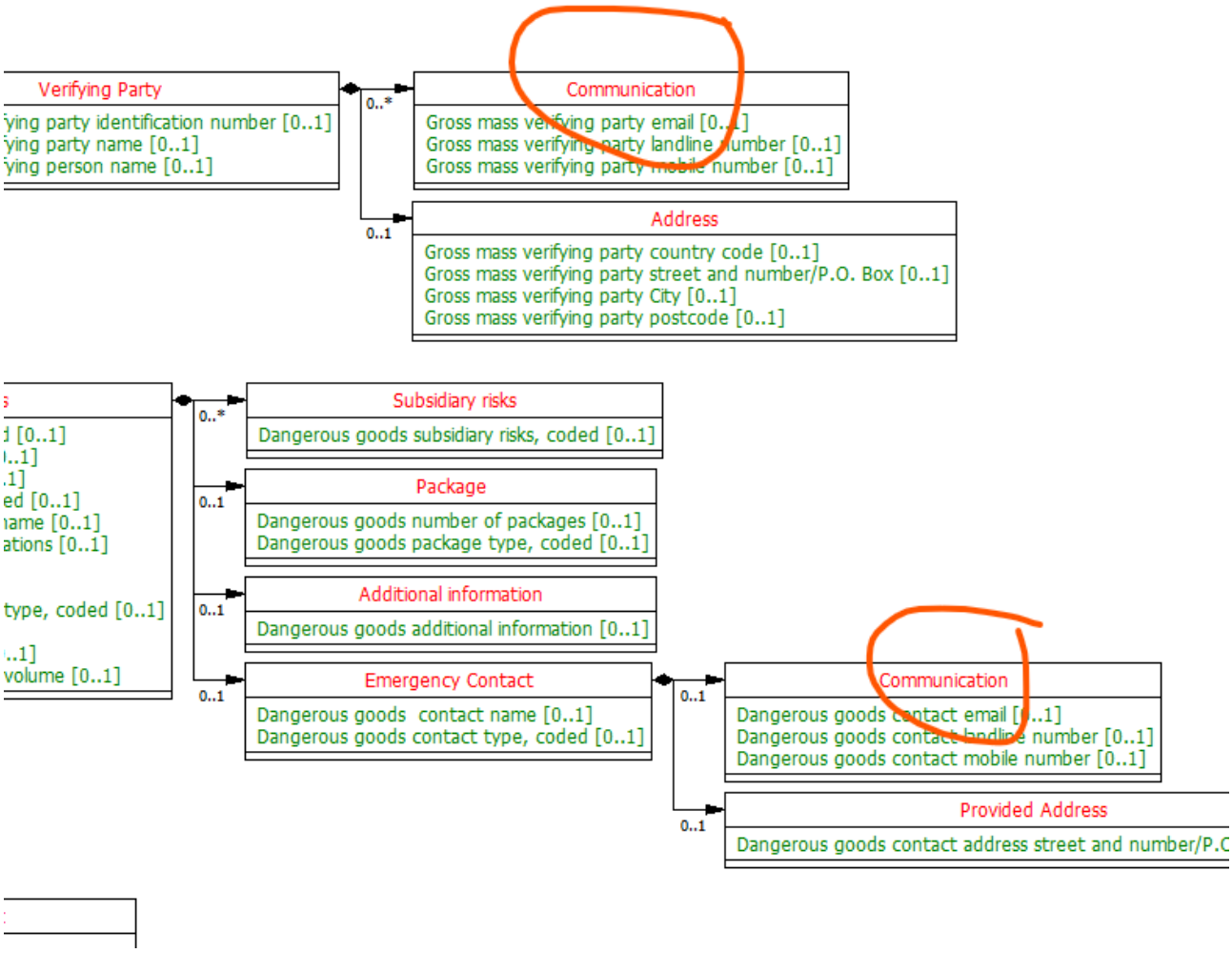


Figure 6 In S-100 object names are unique identifiers for the objects. Therefore duplicate names cannot exist and has to be renamed.

4.3 Step 2 –Implementing the minimal needed (complex) structure

The initial import reflects the structure of the Datasets as a Data Dictionary of needed elements. At least some parts of the structure of the reference model must also be implemented. This need is most evident where more than one instance is expected for a given object. The most apparent examples would be the lists of passengers and crewmembers, assuming that one list would most likely contain multiple persons. In order to identify persons and the attributes related to those persons, additional structure must be adopted.

4.3.1 Passenger list (FAL form 6)

The FAL form 6 , Passenger list is used as the initial Dataset to implement, due to perceived simplicity, and the apparent need to implement multiple instances of passengers within each Dataset.

PASSENGER LIST
(IMO FAL Form 6)

☐ Arrival ☐ Departure ☐ Page Number

1.1 Name of ship			1.2 IMO number			1.3 Call sign							
1.4 Voyage number			2. Port of arrival/departure			3. Date of arrival/departure			4. Flag State of ship				
5. Family name	6. Given names	7. Nationality	8. Date of birth	9. Place of birth	10. Gender	11. Type of identity or travel document	12. Serial number of identity or travel document	13. Issuing State of identity or travel document	14.Expiry date of identity or travel document	15. Port of embarkation	16. Visa number if appropriate	17. Port of disembarkation	18. Transit passenger or not

Figure 7 The FAL 6 form as a conventional Passenger list.

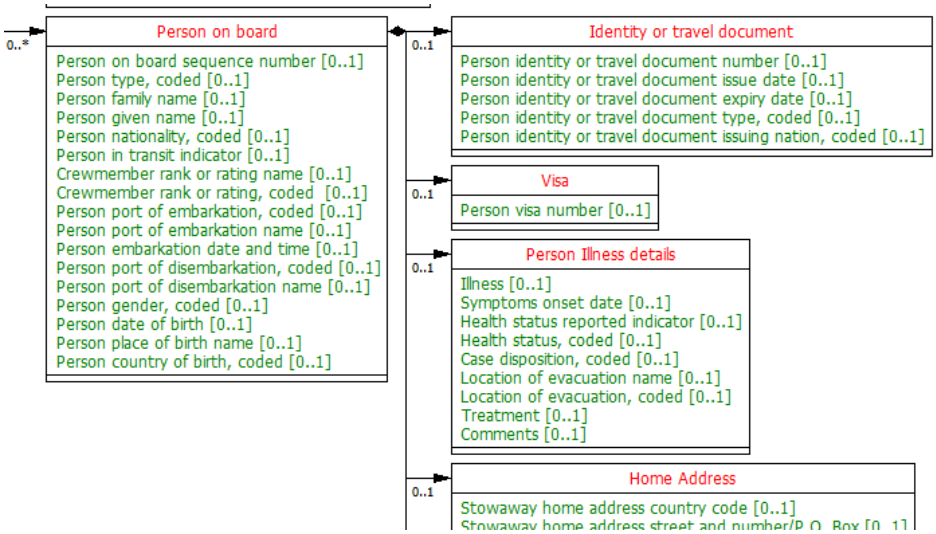


Figure 8 The structure and multiplicities are expressed in the Reference model.

Passenger_List

agent_contact_name	(IMO0002)	text	[0..1]
agent_identification_number	(IMO0007)	text	[0..1]
agent_name	(IMO0010)	text	[0..1]
arrivaldeparture_code	(IMO0013)	text	[0..1]
authentication_date	(IMO0014)	text	[0..1]
authenticator_name	(IMO0016)	text	[0..1]
authenticator_party_identification_number	(IMO0017)	text	[0..1]
date_and_time_of_arrival_actual	(IMO0063)	text	[0..1]
date_and_time_of_arrival_estimated	(IMO0064)	text	[0..1]
date_and_time_of_departure_actual	(IMO0065)	text	[0..1]
date_and_time_of_departure_estimated	(IMO0066)	text	[0..1]
person_in_transit_indicator	(IMO0089)	integer	[0..1]
person_port_of_embarkation_coded	(IMO0091)	text	[0..1]
person_port_of_disembarkation_name	(IMO0092)	text	[0..1]
person_port_of_disembarkation_coded	(IMO0093)	text	[0..1]
person_port_of_embarkation_name	(IMO0094)	text	[0..1]
person_visa_number	(IMO0095)	text	[0..1]
person_country_of_birth_coded	(IMO0096)	text	[0..1]
person_date_of_birth	(IMO0097)	text	[0..1]
person_family_name	(IMO0098)	text	[0..1]
person_gender_coded	(IMO0099)	text	[0..1]
person_given_name	(IMO0100)	text	[0..1]
person_identity_or_travel_document_expiry_date	(IMO0101)	text	[0..1]
person_identity_or_travel_document_issuing_nation_coded	(IMO0102)	text	[0..1]
person_identity_or_travel_document_number	(IMO0103)	text	[0..1]
person_identity_or_travel_document_type_coded	(IMO0104)	text	[0..1]
person_nationality_coded	(IMO0105)	text	[0..1]
person_place_of_birth_name	(IMO0106)	text	[0..1]
person_type_coded	(IMO0107)	text	[0..1]
port_of_arrival_coded	(IMO0108)	text	[0..1]
port_of_arrival_name	(IMO0109)	text	[0..1]

Figure 9 The passenger list might contain several persons. This must be somehow reflected in the implementation.

The Reference model contains a top-down structure, where all elements are expressed as optional. For example, Stowaway address information is available, but is not relevant for a passenger. Thus, the structure must be further restricted on the Dataset- level. The Dataset Passenger list contains Data elements from the Person on Board, Identity or Travel document and Visa only.

In order to allow multiple instances of this information for each passenger, an additional Complex structure has to be implemented. For this implementation, a new complex attribute, with only the filtered set of data, is created and called a *Passenger*. In order to facilitate further data re-use, an additional complex structure called *personOnBoardBasicData* is created. The multiplicity for Passengers is 0..*

Passenger_List

agent_contact_name	(IMO0002)	text	[0..1]
agent_identification_number	(IMO0007)	text	[0..1]
agent_name	(IMO0010)	text	[0..1]
arrivaldeparture_code	(IMO0013)	text	[0..1]
authentication_date	(IMO0014)	text	[0..1]
authenticator_name	(IMO0016)	text	[0..1]
authenticator_party_identification_number	(IMO0017)	text	[0..1]
date_and_time_of_arrival_actual	(IMO0063)	text	[0..1]
date_and_time_of_arrival_estimated	(IMO0064)	text	[0..1]
date_and_time_of_departure_actual	(IMO0065)	text	[0..1]
date_and_time_of_departure_estimated	(IMO0066)	text	[0..1]
passenger	(passenger)		[0..*]
personOnBoardBasicData	(personOnBoardBasicData)		[1..1]
person_on_board_sequence_number	(IMO0044)	integer	[0..1]
person_type_coded	(IMO0107)	text	[0..1]
person_family_name	(IMO0098)	text	[0..1]
person_given_name	(IMO0100)	text	[0..1]
person_nationality_coded	(IMO0105)	text	[0..1]
person_in_transit_indicator	(IMO0089)	integer	[0..1]
crewmember_rank_or_rating_name	(IMO0042)	text	[0..1]
crewmember_rank_or_rating_coded	(IMO0043)	text	[0..1]
person_port_of_embarkation_coded	(IMO0091)	text	[0..1]
person_port_of_embarkation_name	(IMO0094)	text	[0..1]
person_embarkation_date_and_time	(IMO0219)	text	[0..1]
person_port_of_disembarkation_coded	(IMO0093)	text	[0..1]
person_port_of_disembarkation_name	(IMO0092)	text	[0..1]
person_gender_coded	(IMO0099)	text	[0..1]
person_date_of_birth	(IMO0097)	text	[0..1]
person_place_of_birth_name	(IMO0106)	text	[0..1]
person_country_of_birth_coded	(IMO0096)	text	[0..1]
identity_or_travel_document_c	(identity_or_travel_document_c)		[0..1]
person_identity_or_travel_document_number	(IMO0103)	text	[0..1]
person_identity_or_travel_document_issue_date	(IMO0259)	text	[0..1]
person_identity_or_travel_document_expiry_date	(IMO0101)	text	[0..1]
person_identity_or_travel_document_type_coded	(IMO0104)	text	[0..1]
person_identity_or_travel_document_issuing_nation_coded	(IMO0102)	text	[0..1]
visa_c	(visa_c)		[0..1]
person_visa_number	(IMO0095)	text	[0..1]

Figure 10 Passenger list with the adapted structure for allowing multiple passenger data. Elements with a yellow background are those available in the Reference model structure, but not used in the current Dataset. Crewmember- related elements are not relevant to passengers, and these elements should be implemented in a separate container, and used only for Crewmembers. Relevancy of the inconsistent elements related to persons should be evaluated, in order to understand the reason of inconsistency.

The Reference model was found to contain some inconsistent structures compared to the Passenger list dataset. For the purpose of this exercise, the two attributes directly related to crewmembers were removed from the personOnBoardBasicData- structure, but the other attributes missing from the Dataset were kept as presented in the Reference model.

Multiplities are not further modified at this stage, except for the additions of passenger [0,*], allowing for any number of passengers, and the personOnBoardBasicData [1,1] indicating that each passenger must have exactly one (mandatory) element of basic data.

4.3.2 Compare to ISO 28005-3 draft implementation

Comparing the above exercise to the ISO 28005-3 draft shows that a very similar structure was also implemented by ISO. Anyway, as additional structure is created, there is some room for differences within implementations, such as naming of types and low-level structuring of names, addresses etc.

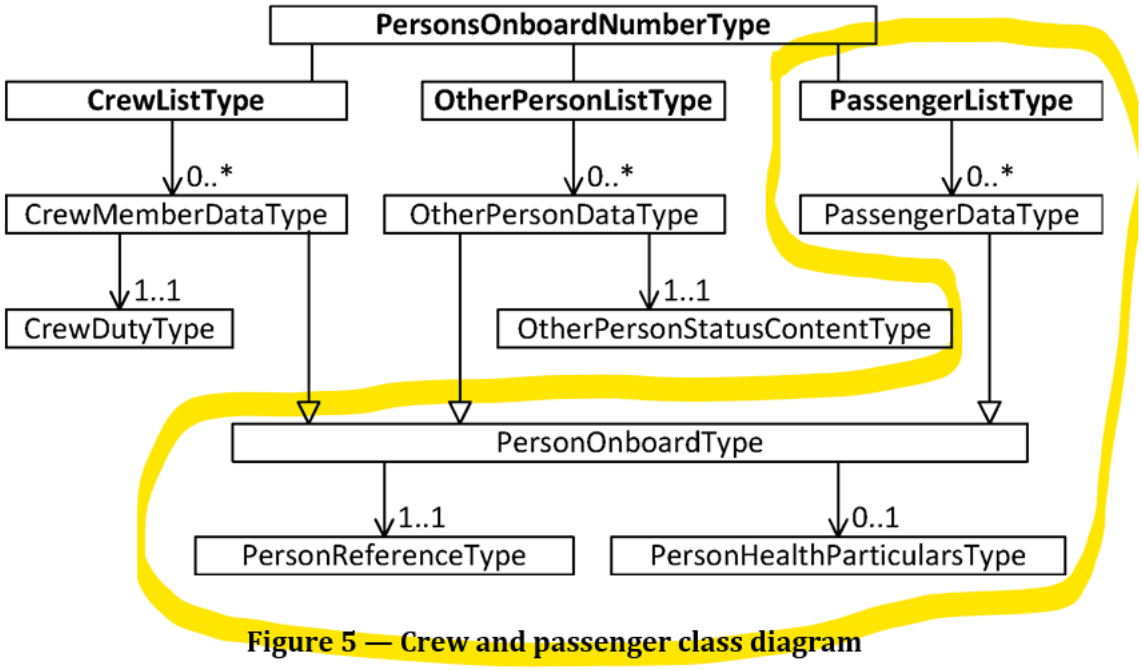


Figure 11 The ISO implementation of the PassengerListType is very similar in structure to the outcome of the excersice. The main diffrence in implementation is the different naming of the object containing data common to all persons onboard (PersonOnboardType vs. personOnBoardBasicData).

4.4 Step 3 – Mapping data format and implementing codelists

The EDIFACT format codes must be translated into primary data types understood by the implementation framework. Codelists must be further restricted to allow only the supported values. Additionally, S-100 requires that enumeration values are defined and provided together with the dataset.

The initial import translated all DataElements into integer, real or text (characterString). Other basic types in S-100 are Date, Time and DateTime, represented as a string conforming to ISO 8601. Predefined derived types include URI, URN and URL.

Complex types are combinations of other types. The passenger- structure described in the previous chapter is an example of a Complex Type.

Continuing with the Passenger list as an example, a few of the Data element formats are considered. Dates and times are converted into the corresponding datatypes, and some of the codelists are implemented.

4.4.1 Implementing codelists

For an S-100 implementation, the items of codelists must be added to the Product specification as an enumeration, or referenced as a codelist by a URL where the data is found. For simplicity of this exercise, codelist-data is added to the data model as enumeration values.

Element *person_identity_or_travel_document_number* can take values Identity card, Registration document or Passport. These options are encoded as an enumeration.

Element *arrival_departure_code*, IMO0013 is an indication of arrival or departure. This element accepts values A = Arrival or D=Departure. This element is implemented as an enumeration with options Arrival or Departure.

Element *person_gender_coded*, IMO0099, is defined as a code representing the gender of the referenced person. UN/EDIFACT Codes 3499 are referenced, and further found to adhere to four codes specified in ISO/IEC 5218, which are:

0 = Not known;

1 = Male;

2 = Female;

9 = Not applicable

Also other codelists such as ISO 3166 country codes etc. should be referenced or implemented, but this need is simply noted in this exercise. Common codelists might also be implemented in the framework and could be referenced out of existing lists.

4.5 Step 4 - Sample S-100 GML dataset (Passenger list)

At this point, the first sample, dataset could be produced. Tools previously created for in-house testing of S-100 datasets were used for this purpose. The sample dataset is conformant to S-100 5.0.0 GML and GML 3.2.2 specifications.

```
<Dataset xmlns="http://www.iho.int/IMO/gml/1.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:gml="http://www.opengis.net/gml/3.2"
 xmlns:S100="http://www.iho.int/s100gml/5.0" xmlns:s100_profile="http://www.iho.int/S-100/profile/s100_gmlProfile"
 xmlns:xlink="http://www.w3.org/1999/xlink" xsi:schemaLocation="http://www.iho.int/IMO/gml/1.0 ../../schemas/IMO/0.0.1/20221121/IMO.xsd"
 gml:id="FIHO.GML.644230ef49e33">
  <gml:boundedBy>
    <gml:Envelope srsDimension="2" srsName="urn:ogc:def:crs:EPSG:4326">
      <gml:lowerCorner>58 19</gml:lowerCorner>
      <gml:upperCorner>61 22</gml:upperCorner>
    </gml:Envelope>
  </gml:boundedBy>
  <S100:DatasetIdentificationInformation>
    <S100:encodingSpecification>S-100 Part 10b</S100:encodingSpecification>
    <S100:encodingSpecificationEdition>1.0</S100:encodingSpecificationEdition>
    <S100:productIdentifier>INT.IHO.IMO.X.X</S100:productIdentifier>
    <S100:productEdition>X.X</S100:productEdition>
    <S100:applicationProfile>1.0</S100:applicationProfile>
    <S100:datasetFileIdentifier>IMO.gml</S100:datasetFileIdentifier>
    <S100:datasetTitle>IMO test dataset, FINLAND</S100:datasetTitle>
    <S100:datasetReferenceDate>2023-04-21</S100:datasetReferenceDate>
    <S100:datasetLanguage>eng</S100:datasetLanguage>
    <S100:datasetAbstract>Test dataset NOT FOR NAVIGATION by stefan.engstrom@traficom.fi, TRAFICOM, FINLAND</S100:datasetAbstract>
    <S100:datasetTopicCategory>transportation</S100:datasetTopicCategory>
    <S100:datasetPurpose>base</S100:datasetPurpose>
    <S100:updateNumber>0</S100:updateNumber>
  </S100:DatasetIdentificationInformation>
  <members>
    <Passenger_List gml:id="fiho.s100.IMO.Passenger_List.0001">
      <agent_contact_name>Bond James</agent_contact_name>
      <agent_identification_number>007</agent_identification_number>
      <agent_name>M5</agent_name>
      <arrivaldeparture_code code="1">Arrival</arrivaldeparture_code>
      <authentication_date>2022-12-30</authentication_date>
      <authenticator_name>Moneypenny Miss</authenticator_name>
      <date_and_time_of_arrival_estimated>2023-01-01T00:00Z</date_and_time_of_arrival_estimated>
      <passenger>
        <personOnBoardBasicData>
          <person_on_board_sequence_number>1</person_on_board_sequence_number>
          <person_family_name>Khan</person_family_name>
          <person_given_name>Kamal</person_given_name>
          <person_port_of_embarkation_name>Udaipur</person_port_of_embarkation_name>
          <person_country_of_birth_coded>FI</person_country_of_birth_coded>
        </personOnBoardBasicData>
        <identity_or_travel_document_c>
          <person_identity_or_travel_document_issue_date>1970-01-01</person_identity_or_travel_document_issue_date>
          <person_identity_or_travel_document_type_coded code="36">Identity card</person_identity_or_travel_document_type_coded>
        </identity_or_travel_document_c>
      </passenger>
      <passenger>
        <personOnBoardBasicData>
          <person_on_board_sequence_number>2</person_on_board_sequence_number>
          <person_family_name>Powers</person_family_name>
          <person_given_name>Max</person_given_name>
          <person_port_of_embarkation_name>London</person_port_of_embarkation_name>
        </personOnBoardBasicData>
      </passenger>
      <port_of_arrival_name>Helsinki</port_of_arrival_name>
      <ship_call_sign>HELP</ship_call_sign>
      <ship_name>Agents of the Seas</ship_name>
    </Passenger_List>
  </members>
</Dataset>
```

Figure 12 A Passenger list of imaginary vessel "Agents of the Seas" in S-100 GML format, with 2 passengers on board. The structure is adhering to the Feature Catalogue as imported with the above described amendments. As most elements are optional, only selected data is included.

4.6 Step 5 - Comparing specific data elements, identifiers and structures

This part of the exercise investigates and compares structures and elements of special interest for harmonization of S-100 and the IMO Compendium. The Just in time dataset is selected for further investigation, as JIT- data;

- refers to features also present in IHO S-101 (ENC) and S-131 (Harbour infrastructure)
- contains geographical and named location data

The JIT dataset is still under development in FAL 46 version of IMO Compendium, and therefore no attempt was made at this point to further implement this Dataset.

Just_In_Time_Concept

authentication_date	(IMO0014)	Date	[0..1]
authenticator_name	(IMO0016)	text	[0..1]
authenticator_party_identification_number	(IMO0017)	text	[0..1]
date_and_time_of_arrival_actual	(IMO0063)	DateTime	[0..1]
date_and_time_of_arrival_estimated	(IMO0064)	DateTime	[0..1]
date_and_time_of_departure_actual	(IMO0065)	DateTime	[0..1]
date_and_time_of_departure_estimated	(IMO0066)	DateTime	[0..1]
port_of_arrival_coded	(IMO0108)	text	[0..1]
authenticator_role_coded	(IMO0128)	text	[0..1]
ship_stay_reference_number	(IMO0153)	text	[0..1]
port_facility_coded	(IMO0184)	text	[0..1]
port_facility_name	(IMO0185)	text	[0..1]
anchorage_name	(IMO0229)	text	[0..1]
terminal_name	(IMO0230)	text	[0..1]
pilot_boarding_place	(IMO0231)	text	[0..1]
berth_name	(IMO0232)	text	[0..1]
berth_position	(IMO0233)	text	[0..1]
date_and_time_of_arrival_requested	(IMO0234)	DateTime	[0..1]
date_and_time_of_arrival_planned	(IMO0235)	DateTime	[0..1]
date_and_time_of_departure_requested	(IMO0236)	DateTime	[0..1]
date_and_time_of_departure_planned	(IMO0237)	DateTime	[0..1]
message_function_code	(IMO0305)	text	[0..1]
trade_service_identifier	(IMO0536)	text	[0..1]
distance_to_destination	(IMO0537)	integer	[0..1]
average_speed	(IMO0538)	real	[0..1]
date_and_time_to_location_in_port_actual	(IMO0540)	DateTime	[0..1]
date_and_time_to_location_in_port_estimated	(IMO0541)	DateTime	[0..1]
date_and_time_to_location_in_port_requested	(IMO0542)	DateTime	[0..1]
date_and_time_to_location_in_port_planned	(IMO0543)	DateTime	[0..1]
location_in_port_latitude	(IMO0544)	text	[0..1]
location_in_port_longitude	(IMO0545)	text	[0..1]
anchorage_coded	(IMO0546)	To be defined	[0..1]
terminal_coded	(IMO0547)	To be defined	[0..1]
berth_coded	(IMO0548)	To be defined	[0..1]

Figure 13 JIT- dataset

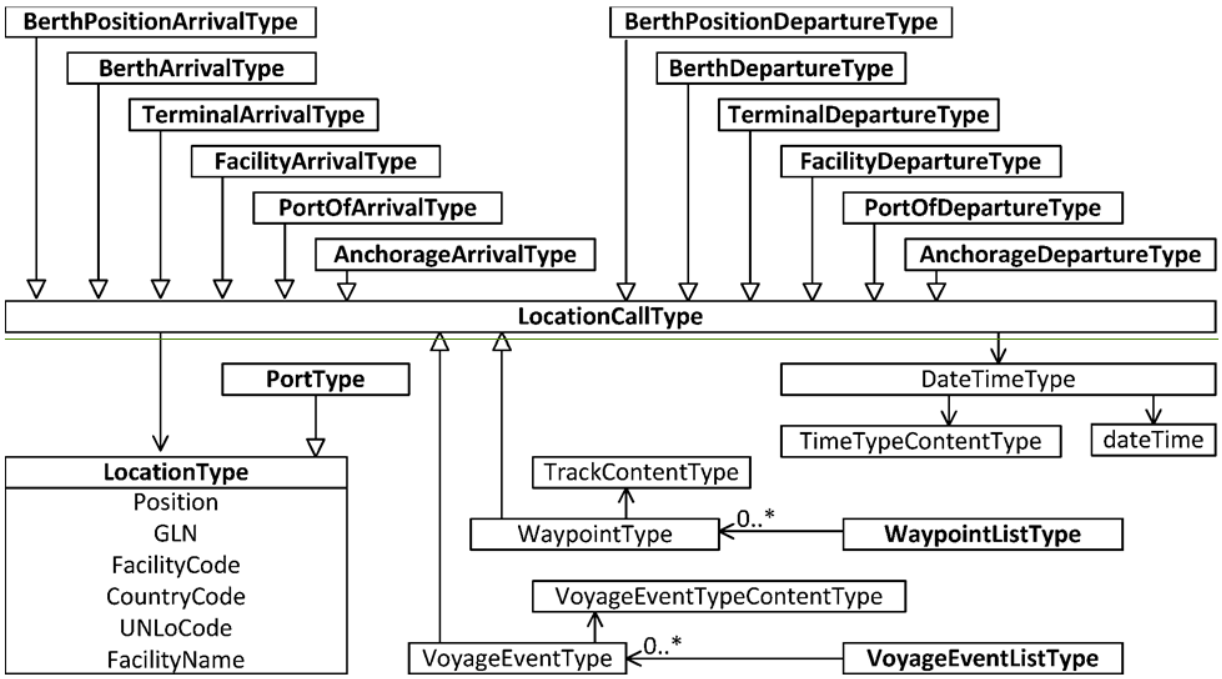


Figure 14 ISO standard representation of types containing location.

4.6.1 Geographical location

The IMO compendium refers to ISO 6709, Standard representation of geographic point location by coordinates. All locations are 2 dimensional points in the, represented as thematic attributes (latitude and longitude).

S-100 / ISO-19100 does not use thematic attributes for locations. Locations are represented by spatial objects, connected to the features. Each S-100 FeatureType can have an associated location expressed by a point, line or surface geometry. In practice, S-100 / ISO-19100 requires each position specified as latitude / longitude to be modelled as a separate feature with a point geometry. This is similar to the implementation in the ISO280005 standard draft, where a “LocationType” is defined. In an S-100 implementation, each type utilizing the LocationType would be a Feature.

4.6.2 Object identifiers and mandatory attributes

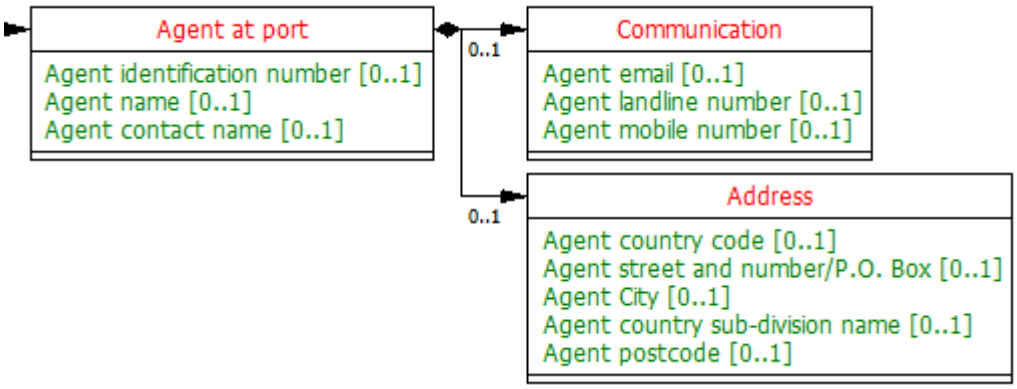
Objects in Figure 13 "LocationType" contains a list of possible identifiers to use for identification of features. It is important that similar identifiers are available across products. Also, an open list with several options for identification using either GLN, FacilityName, FacilityCode or UNLocode might lead to situations, where the same feature is given different identifiers. Especially for identifiers, it might be feasible to add mandatory or recommended identifiers.

4.6.3 Objects and data encapsulation

IMO datasets are described as data element dictionaries, without structure. The complex structure is separately described in the IMO reference model. Implementation of complex structure requires cross-referencing both.

Implementation of the passenger list indicated, that multiple passengers in one list had to be implemented using a complex structure. Complex structure is separately obtained from the reference model, but has additionally to be filtered according to the needs of the current dataset. In S-100, each such filtered version of a complex structure would become a new complex structure. Therefore, the S-100 approach would probably be to split such structures into objects, based on the needs of the Datasets.

As an example, *Agent at port* data is used in many datasets, but the extent of data is different. Some Datasets use only basic agent data (Agent at port) and some contain more detailed data (also Communication & Address). This could be reflected in the reference structure and be split into separate structures;



- agentBasicData
  - Agent identification number
  - Agent name
  - Agent contact name
- agentExtendedData
  - Agent email
  - Agent landline number
  - Agent mobile number
  - Agent country code
  - Agent street and number
  - Agent city
  - Agent country sub-division name
  - Agent postcode

A S-100 style Passenger List would reference the complex structure *agentBasicData* instead of the individual attributers, but General Declaration would reference also the *agentExtendedData*.

Passenger List

Data number	Data element
IMO0002	Agent contact name
IMO0007	Agent identification number
IMO0010	Agent name

Figure 15 An S-100 implementation of Passenger list would reference agentBasicData instead of the individual elements.

General Declaration

Data number	Data element	Defini
IMO0001	Agent City	The ci
IMO0002	Agent contact name	The ag
IMO0003	Agent country code	A code
IMO0004	Agent country sub-division name	The co
IMO0006	Agent email	The em
IMO0007	Agent identification number	The re
IMO0008	Agent landline number	The la
IMO0009	Agent mobile number	The m
IMO0010	Agent name	The na
IMO0011	Agent postcode	An alp
IMO0012	Agent street and number/P.O. Box	The po
IMO0013	Arrival/departure code	A code

Figure 16 The General Declaration could additionally reference the extended data (which is practically the contact data)

S-100 consider attributes encapsulated. Encapsulation means, that the object "owns" the attributes and attributes must always be accessed through the object. Even though the IMO Compendium Data elements have a position within a container in the reference data model, each Data element is still "self contained", such that the attribute can exist also outside of the container. As an example the structure of Agent data in S-100 might be as follows.

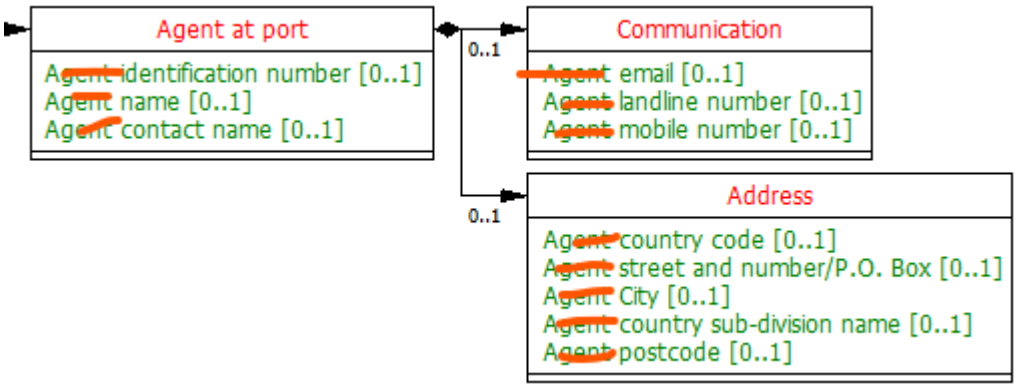


Figure 17 Encapsulation gives attributes context by position within a container. Thus, an S-100 style implementation would most likely omit the pefix 'Agent_' within the container.

- agentBasicData
  - Agent at Port
    - identification number
    - name
    - contact name
- agentExtendedData
  - Communication
    - email

        - landline number
        - mobile number
    - Address
        - country code
        - street and number
        - city
        - country sub-division name
        - postcode

S-100 encapsulated attributes would not repeat the prefix 'Agent_' as the context of an encapsulated attribute is determined by the container. This also means, that referenced attributes can be defined in a more general way.

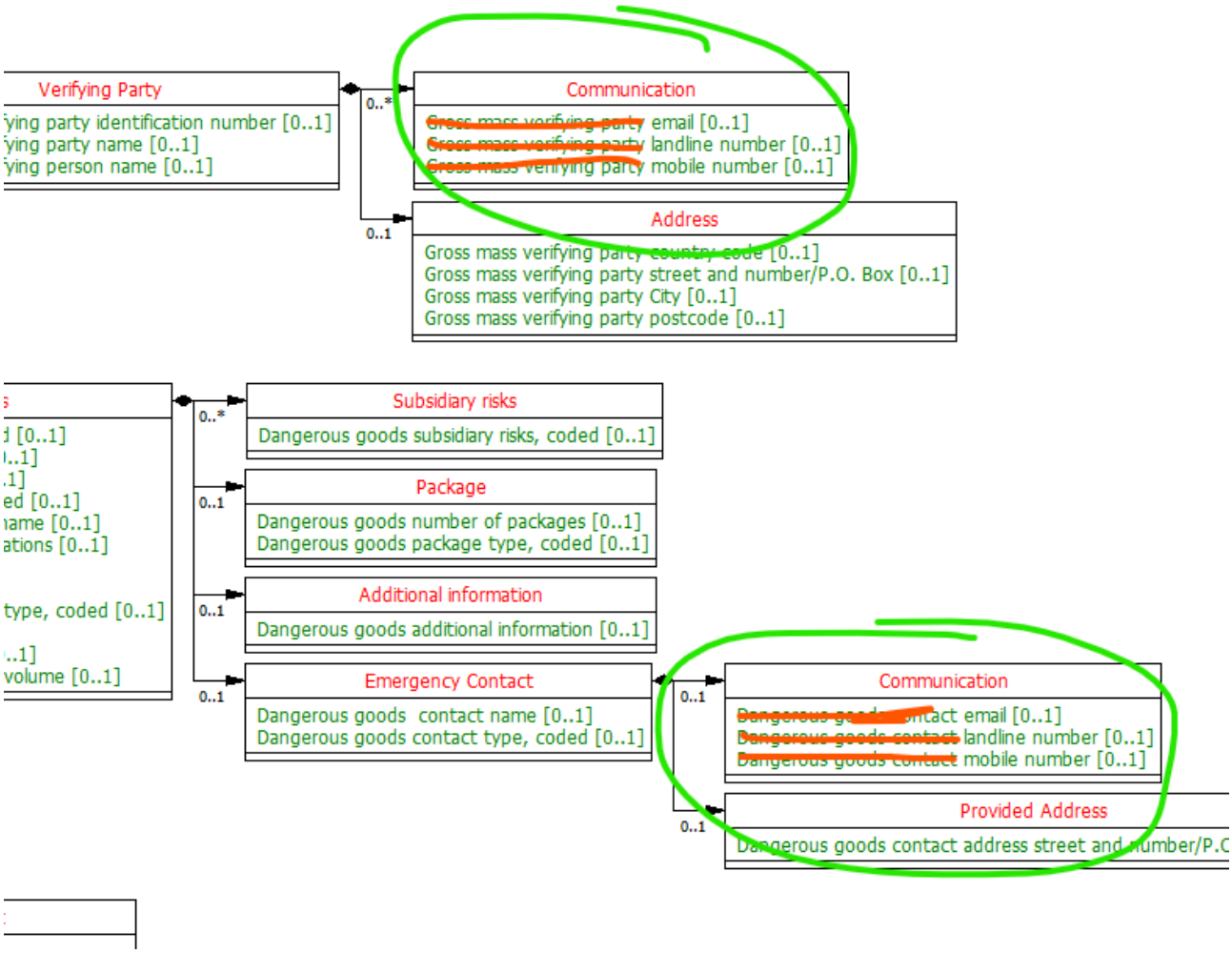


Figure 18 With encapsulation, the Communication- objects of the other instances could become similar to the agentExtendedData::Communication , and these could all refer to the same object.