

## Paper for Consideration by NIPWG

## Report on creation of Canadian S-123 datasets

<b>Submitted by:</b>	Canadian Coast Guard
<b>Executive Summary:</b>	This paper contains discussions and proposals for feedback to NIPWG for improving the S-123 Product Specification based on the experiences gained from making S-123 Datasets from the Canadian Coastguard Radio Aids to Marine Navigation (RAMN) document.
<b>Related Documents:</b>	S-123 v.1.0.0 Report on Feedback to S-123 edition 1.0.0
<b>Related Projects:</b>	S-123

**Introduction / Background**

The Canadian Coast Guard (CCG) has contracted Caris to create the Canadian coverage of S-123 data. The source document is the official Canadian publication Radio Aids to Marine Navigation (RAMN).

**Analysis/Discussion**

The contract was awarded in two phases; phase 1 was to test the capabilities and see what the results would be. The Great Lakes was chosen as the test area due to its variation of types of data to capture and it's relatively easily definable geographic area. Along with the data creation, the contract included a detailed report on the exercise as to help CCG understand the process better and the state of S-123.

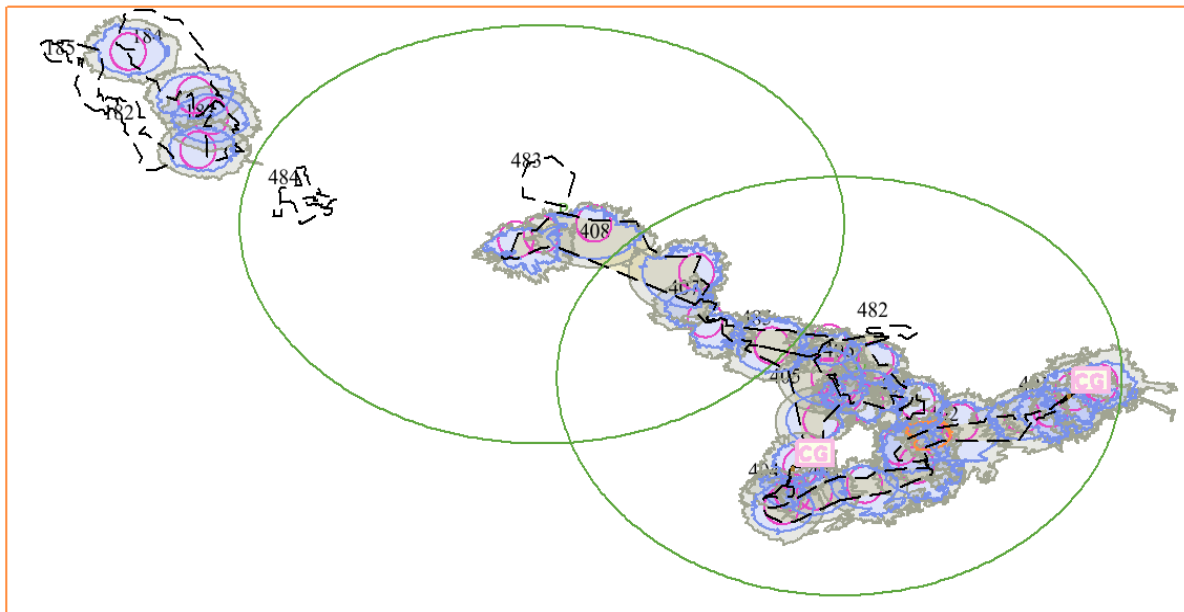


Figure 1 - Radio Services in Great Lakes region of Canada

The exercise proved very helpful in identifying issues with the IHO S-123 specifications. Following are some recommendations from Caris for improvement to the specification based on their experience creating a test dataset for S-123. The exercise has also highlighted some issues at the S-100 level for which Caris has recommended some solutions.

## 1. S-123 main specification

### 1.1. Model explanations

Section 6.2 Application schema overlaps or duplicates content which is in the separate Application schema document, DCEG and Feature Catalogue. Concerns related to this are:

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

- duplication of effort for maintenance
- potential for inconsistencies
- difficult to use as a reference because content is scattered among multiple documents

### 1.1.1. Recommendation:

Use the product specification to define the overarching intent and structure decisions of the product as well as the profile of what content and options from S-100 are used in the product. Move the details of the content model from section 6 to the encoding guide. Section 6 can include key patterns/examples and refer to the application schema and encoding guide documents for details.

## 1.2. Model Diagrams and Tables

The UML diagrams, table of attributes and feature catalogue (FC) are different presentations of the same content. A production system would typically use the feature catalogue or a defined schema to offer the encoder the available choices and options for creating features, entering attributes and relationships. The UML diagrams are likely most valued by the modelers who are defining the product contents. The tables describing the UML diagrams are duplicated in the Application schema and the encoding guide.

### 1.2.1. Recommendation:

Consider including the UML diagrams, tabular representation and encoding nuances all in the encoding guide. This would save users from needing to jump back and forth between documents when trying to understand how to encode a given feature such as a Radio Station. In the product spec only keep top level UML diagrams that indicate how the product spec is profiled from S-100.

There may be value in having a UML representation available as a separate document and/or as a machine-readable format and it needs to be kept in sync with the S-100 FC.

## 1.3. Dataset Naming Convention

The dataset naming convention, described in Section 11.6 of the S-123 Product specification is being harmonized across S-100 product specs. Ideally this would be defined as part of S-100 proper but currently the S-97 specification provides guidance meant to address this issue.

Producer and user systems will both benefit from consistent file naming and exchange set structure patterns:

- Consistency across products
- Consistency for solution providers and testers, not requiring product specific implementations
- Better for data discovery and searches for products/datasets
- Better for organizing/classifying/reporting/discovering data

### 1.3.1. Recommendation:

Recommend aligning dataset naming convention with S-97 section 5.2.21 and other product specs;

XXXXXXXXXXXXXXXXXXXX

XXX is the product code, e.g. 123 is for Maritime Radio Service.

YYYY is the producer code according to the Producer Code Register. Padded with '0' for agency codes less than 4 characters.

0000 is an arbitrary length unique code in alphanumeric characters.

## 1.4. Catalogue file name convention

Section 11.8 of the product spec shows a catalogue xml file name unique to S-123 as 'CATALOG.123.XML'.

Since the preparation of this spec efforts have been made to allow a single exchange set to hold multiple products. It is also simpler for production, testing and end user systems if consistent patterns are used.

#### 1.4.1. Recommendation:

Use a common filename catalog.xml file for consistent implementations and discovery.

### 1.6. Appendix D-1 GML Data Format Overview

Seems to overlap or duplicate the content in the separate GML Data Format document.

#### 1.6.1. Recommendation:

The main product spec should include information about encoding choices and patterns but the detailed GML schema definition could be a separate document.

## 2. Application Schema, Data Classification and Encoding Guide (DCEG) and Feature Catalogue (FC)

### 2.1. Overlap of content

As discussed above in the section on the S-123 Main Specification, the duplication of content in various representations is a double-edged sword. It allows users to choose the presentation they are most comfortable with but presents difficulties in maintenance and consistency for producers and users.

### 2.2. Names/Codes

Consistently include both the name and CamelCase code in the sections describing each type. Readers should be able to search by 'Radio Station' or 'RadioStation'. This seems to have been done in most cases but there are a few exceptions such as **NavigationalMeteorologicalArea**. In DCEG section 5.7 it is misspelled as 'NavigationalMeterorologicalArea' and the name is 'NAVAREA/METAREA'.

Changed case of Information association codes to UpperCamelCase for consistency.

Information type **AbstractRxN** is missing a role for the information association binding **InclusionType** to Applicability, assuming to use **isApplicableTo**.

#### 2.2.1. Recommendations:

- Review Names and codes for consistency. Adjust DCEG section 5.7 NavigationalMeteorologicalArea and name.
- Information types should be UpperCamelCase. XML FC and other docs.
- Fix missing InclusionType for Information type AbstractRxN. XML FC

### 2.3. Discussion about modeling considerations

#### 2.3.1. Feature Type

Abstract base class for features.

Inconsistency between DCEG xml FC.

DCEG section 5.1 shows **textContent** 0,\* but FC is 0,1

#### 2.3.2. Coast Guard Station

In the RAMN document the regional radio stations and services are organized according to the regional Coast Guard Stations from which they are managed. Coast Guard Stations remotely control and provide the services offered by a set of Radio Stations. Contact information and hours of service are specific to the Coast Guard Station.

S-123 does not provide a means to encode the relationships between Radio Station and Coast Guard Station. The only feature link available for **CoastguardStation** is to a **TextPlacement**. Machine readable navigation between the **CoastguardStation** and related **RadioStation** features seems to be through a link to a common **ContactDetails** information type. It seems more efficient in this context to be able to navigate directly between **CoastguardStation** and **RadioStation**.

The contact details are defined for the Coast Guard station and not individual Radio Stations.

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

### 2.3.10. RadioStation

The Radio Station is primarily the location of the antenna/equipment for radio receiving/transmitting.

Most of the details of the services provided by the radio station are encoded in the **RadioServiceArea** features which provide coverage and details about services which are provided by the radio station. The same radio station could offer different service coverage areas due to frequency and power etc.

In the context of the RAMN content Radio stations are unmanned and remotely operated from a common coast Guard station. There should be an association/relation to the entity which controls the **RadioStation** such as a coast Guard station.

**RadioStation** should also have an attribute to indicate that it is remotely operated or manned.

The S-123 specs state that only four of the sub-attributes of radiocommunications complex attribute are applicable to **RadioStation**. In the S-123 Feature Catalogue, **RadioStation** is bound to the radiocommunications. Since the S-100 FC model does not support restriction of sub-attributes on a bound complex attribute it would be better to make a different complex attribute with only the four desired sub-attributes.

This may have been considered and partially implemented as there is such a complex attribute **radioStationCommunicationDescription** already in the FC which is unused, and it seems that this should be bound to **RadioStation** instead of the **radiocommunications** attribute.

It is far better to include some attributes into different complex containers for specific purposes than to have textual notes in specification documents which are not machine readable/usable to restrict the content. Using this separate complex dedicated to **RadioStation** would avoid attributes being populated by mistake and then having to run separate validation checks to find/fix the encoding mistakes after the fact.

If the communications channel, frequencies etc are populated on the **RadioServiceArea** features associated with a **RadioStation** then it would be redundant to encode channel and frequency on the **RadioStation**. In DCEG remarks under 5.10.1 it states:

*EACH VHF-CHANNEL AND COMMUNICATIONS FREQUENCY SHOULD BE INDICATED, USING THE SUB-ATTRIBUTES COMMUNICATION CHANNEL OR SIGNALFREQUENCY OF COMPLEX ATTRIBUTE RADIOCOMMUNICATIONS*

The DCEG should be updated to indicate that **RadioStation** communication attributes are not necessary if encoded in related features.

#### 2.3.2.1. Recommendations:

- **CoastguardStation** should have a relationship with **RadioStation** (manages,managedBy) or (controls, controlled by). If this relationship was available, it would have the following benefits:
  - It would not be necessary to link **RadioStation** to **ContactDetails** and **ServiceHours** as this would be established through the **CoastguardStation** unless specific to a site. Thus, reducing the number of links to manage.
  - If the **RadioStation** had an attribute indicating it is remotely controlled, then the link to **CoastguardStation** would provide the details.
  - It would provide a more direct linkage to navigate between **RadioStation** features and a **CoastguardStation**. This would be used by user query or to produce documents like RAMN from the digital S-123 data.
- Use a separate complex attribute such as **radioStationCommunicationDescription** to carry only the attributes intended for **RadioStation** instead of needing special notes and validations.
- Change DCEG to indicate **RadioStation** communication information only needed if not available through associated **RadioServiceArea** features.

#### 2.3.10.1. Recommendations:

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

- 
- 
- 

### 2.3.3. GMDSSAreas

In S-123 **GMDSSArea** is defined as having a Surface geometry that can be associated to one or more **RadioStation** features. The association uses the role **serviceArea**. However, it is the **RadioServiceArea** features that define the service areas and a **RadioStation** can support multiple types of services and service areas.

A **RadioStation** can thus be associated with multiple GMDSS areas but the actual coverage is represented by specific **RadioServiceArea** features.

It seems logical that a **GMDSSArea** feature would be better associated with **RadioServiceArea** or **RadioServiceAreaAggregate** features.

Benefits of allowing **GMDSSArea** to be associated with **RadioServiceArea** or **RadioServiceAreaAggregate**:

- Direct navigation between the **GMDSSArea** and the actual service coverages.
- **GMDSSArea** surface could be optional as the actual coverage could be defined by a collection/aggregation of associated **RadioServiceArea** features. This would also reduce the size of the dataset and complexity of managing/maintaining complex surface polygons.
- Maintenance of **GMDSSArea** would only be to add or remove relationships to **RadioServiceArea** or **RadioServiceAreaAggregate** features instead of maintaining complex polygons.

#### 2.3.3.1. Recommendations:

- Associate **GMDSSArea** with **RadioServiceArea** or **RadioServiceAreaAggregate** with role **serviceArea**
- Allow **GMDSSArea** primitive type to be None when associated with Radio Service Area features.

### 2.3.4. InmarsatOceanRegionArea

Seems to overlap with GMDSSArea A3 zones. There doesn't seem to be much attribution or description about what these areas are providing.

### 2.3.5. Landmarks

Landmark features are used only if needed to encode a location relevant to **radiocommunications** but for which a radio service or station is not appropriate. The related radio communications information must be encoded using a **RadioStation** and/or **RadioServiceArea**.

It seems there would be an implicit association between a **Landmark** and a **RadioStation** based on a shared geometry, however, it might be useful to allow a feature association to indicate they are not just collocated.

#### 2.3.5.1. Recommendation:

- Associate **Landmark** with **RadioStation** or **RadioServiceArea**.

### 2.3.6. NavigationalMeteorologicalArea

This defines the geographic areas in which various governments are responsible for navigation and weather warnings.

These have optional associations with **RadioStation** features but it might make more sense to associate with **RadioServiceAreas** (and Broadcasts, if they were an information type) which are providing navigational and weather warnings.

Perhaps explicit links to either Radio Stations or Radio Service Areas would be redundant if the solution is just any Radio Service areas which intersect with the **NavigationalMeteorologicalArea**.

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

The remark under DCEG 5.7 indicates: NAVTEX transmitting station identification characters are allocated within the same areas. Perhaps it makes sense to have a formal association between **NavigationalMeteorologicalArea** and **NavtexStationArea**.

Perhaps there should be a **ContactDetails** for NAVAREAs, but it looks like S-123 does not allow for that. RAMN 4.3.3.1 says:

*Comments concerning the reception of NAVAREA XVII and XVIII broadcasts, especially above 75°N, would be appreciated and should be sent to: NAVAREA XVII and XVIII Prescott MCTS Centre Telephone: 613-925-0666 Facsimile: 613-925-4519 E-mail: navarea17.18@innav.gc.ca*

#### 2.3.6.1. Recommendations:

- Fix the spelling mistake of **NavigationalMeteorologicalArea** in table under 5.7 of DCEG.
- Consider associating **NavigationalMeteorologicalArea** with **RadioServiceArea** or **RadioServiceAreaAggregate** features.
- Consider association between **NavigationalMeteorologicalArea** and **NavtexStationArea**.
- Consider adding possible association between **NavigationalMeteorologicalArea** and **ContactDetails**.

#### 2.3.7. NavtexStationArea

The **NavtexStationArea** appears to be an administration area within which NAVTEX coverages are defined by way of **RadioServiceArea** features.

The **NavtexStationArea** does not carry the **radiocommunications** complex attribute which seems to be the only place where broadcast details can be provided so it would appear that the region would need to be overlapped using a **RadioServiceArea** having radiocommunications with a **categoryOfRadioMethods=6(NAVTEX)** and details about times and content of the broadcasts.

If broadcasts were modeled as an information type then it could also be associated with the **NavtexStationArea** and the **RadioServiceArea**.

Perhaps it would be useful to have the **NavtexStationArea** associated with the **RadioServiceArea(s)** which describe the actual coverage and provide the service/broadcast information.

In Canada there are NAVTEX where the **txIdentChar** is different for English or French yet the **txIdentChar** has a multiplicity of 1. This means the **NavtexStationArea** would need to be duplicated where two language codes are used.

#### 2.3.7.1. Recommendations:

- Consider an association between **NavtexStationArea** and **RadioServiceArea**.
- Allow a **NavtexStationArea** to have multiple **txIdentChar** attributes or make a complex that pairs language and **txIdentChar**

#### 2.3.8. RadioServiceArea

**RadioServiceArea** seems to be one of the most commonly encoded features in S-123 in order to provide details of each service content and coverage.

It might be useful to consider breaking this into sub-types in order to more distinctly differentiate between types of service such as VHF, MF, AMVER, NAVTEX etc. This would also make it easier to handle portrayal and conversion to formatted documents.

In RAMN, sometimes the same broadcast is available from multiple **RadioStation** sites, remote controlled from one coast Guard station which is broadcasting the content.

Consider splitting **radioCommunications** attribute into communications details vs. broadcast details to better define which attributes are applicable. Model these as information types so that the common information can be encoded/maintained once and shared by all the **RadioServiceArea** features providing the broadcast service and  
Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

to other features such as **WeatherForecastWarningArea** for which the content is about. That way a location search would find the overlapping service area and be able to list the broadcasts, or a search for a broadcast could list the service areas and referenced locations.

Note: DCEG for **RadioServiceArea** identifies that the Service can have an explicit Surface geometry or can be an aggregation of **RadioStation** features. In the xml FC the permitted Primitive of noGeometry has not been included. The permitted primitives for **RadioServiceArea** should be as follows:

```
<S100FC:permittedPrimitives>surface</S100FC:permittedPrimitives>
```

```
<S100FC:permittedPrimitives>noGeometry</S100FC:permittedPrimitives>
```

#### 2.3.8.1. Recommendations:

- Consider breaking **RadioServiceArea** into more specialized types
- Consider splitting **radiocommunications** to separate broadcast details
- Consider modeling broadcast details as information types to associate with multiple services and regions where the broadcast is available or regions that the broadcasts concern.
- Consider other recommendations related to **radiocommunications** in a dedicated section of this document.
- Error in xml FC.

#### 2.3.9. Fuzzy Areas or RadioServiceAreaAggregate

##### 2.3.9.1. Encoding confusion

In the Canadian region there are contours available for 90% coverage and 50% coverage. There was some confusion about how to assemble the Aggregate.

The 90% contour fits inside of the 50% contour.

##### Questions arising.

Are Indeterminate zones needed for both 90% and 50% coverage regions?

Does the RadioServiceArea used as a core service area correspond to the 90% coverage contour or the sum/outer limit of the 50% contour.

In DCEG section 6 it states:

*A geographic feature other than Indeterminate Zone is required as the core feature. The totality of the area where the information may apply consists of the union of the location geometries encoded in the core and indeterminate zone features.*

and

*The geometry for the core feature should correspond to a very high confidence level (sufficient to encode it in an ENC if it were an ENC dataset).*

The decision then was to capture the 90% contour as the core RadioServiceArea and make a surrounding Indeterminate area for the 50% coverage and then aggregate these into a RadioServiceAreaAggregate along with the central RadioStation.

Which still leaves the question of whether an indeterminate area is needed under the RadioServiceArea?

##### 2.3.9.2. Documentation inconsistency

In DCEG section 6.2.1.2 Statistical Confidence, it describes the enumerations 'Virtually certain', 'High likelihood', 'Medium likelihood' and 'Low likelihood' and how the percentages are mapped to these enumerations. These enumerations are repeated again in section 6.2.1.3 but without the percentage distinctions. The attribute for these enumerations is called **informationConfidence**. In the **informationConfidence** definition of the Feature Catalogue the percentage distinctions is not included. It is the FC definitions/remarks that are most likely to be presented to the user during digitizing/editing of the content.

### 2.3.9.3. Recommendations:

- Improve DCEG encoding guidance related to core features and indeterminate zones with worked examples
- Resolve inconsistencies and improve description of **informationConfidence** attribute.

### 2.3.11. Weather forecast/warning area

**WeatherForecastWarningArea** are understood to be identifiable regions for which forecasts are available. Ideally these weather areas would then need to be linked to the **RadioServiceArea** which are broadcasting the information about them, but the model does not include this association. The model only includes an association to **RadioStation** but since a **RadioStation** can have different service areas for different purposes it may not be clear from a machine-readable perspective which service areas provide info about the weather region.

The text of the service area may list the number identifiers of the weather area such as "Weather forecast and wave height forecast for marine forecast areas 215, 217 and 219." But it would be better if there was a more direct link between the broadcast service and the areas it provides information about. The only place to put these numeric identifiers on the **WeatherForecastWarningArea** is the **featureName**. Perhaps a separate field for identifier would be useful.

If the broadcasts were Information Types, then the **WeatherForecastWarningArea** could have a link to the broadcasts which give information about the region. Then the **RadioServiceArea** features which are also associated to those broadcasts could be determined.

Forecast areas are geographical features describing regions for which forecasts are available (request or broadcasts) and there may be multiple **RadioServiceArea** features which broadcast information about areas which do not physically overlap the Radio coverage. For example, provision of weather information for a place further along the route. Currently the **WeatherForecastWarningArea** includes a relationship to the **RadioStation** but it should include a relationship to the **RadioServiceArea** which is providing the information (since a **RadioStation** can offer multiple service areas which provide info about neighbouring regions). Also, if Broadcasts were modeled as information types the actual broadcasts which include information about a forecast area could be easily identified by machine.

There are some categories of forecast and warning areas in RAMN that don't seem to have a match with **categoryOfFrcstAndWarningArea** enumerations, such as 'Inland Waters', 'Commercial Shipping Waters' and 'Major Inland Waters'.

#### 2.3.11.1. Recommendations:

- Allow an association between **WeatherForecastWarningArea** and **RadioServiceArea**.
- Improve the model to allow unique identifiers (such as numeric or coded strings) instead of **featureName**.
- Consider modeling broadcasts as information types to be shared between multiple **RadioServiceArea** features and associated to the forecast areas the broadcasts are about. This would support queries from a location to associated broadcasts and vice versa.
- Consider whether additional **categoryOfFrcstAndWarningArea** enumeration values are warranted.

### 2.3.12. radiocommunications

The **radiocommunications** complex attribute is indeed complex. It seems to be modeled as a sort of one size fits all structure but in any given instance only a portion of the attributes are populated or relevant.

#### 2.3.12.1. Broadcasts

It seems that broadcasts could be modeled with a dedicated structure, a separate complex attribute tailored specifically for broadcasts.

Consider modeling broadcast details as an information type using only the attributes of radiocommunications that are related to modeling broadcasts. This would allow broadcast details to be shared/linked between **RadioServiceArea(s)** and **NavtexStationArea(s)** and **WeatherForecastWarningArea(s)**.

#### 2.3.12.2. Languages of services



Services such as broadcasts can be in different languages. The **RadioServiceArea** has a **languageInformation** attribute but that seems more for general information about services offered. It would be useful to have a machine-readable language indication attribute within the **radiocommunications** complex to indicate the language of a service or broadcast. Separate **radiocommunications** complex instances would be created for services where languages are broadcast/supported on specified channels or frequencies. Language information can be included as text within the **transmissionContent** attribute but that does not allow machine readability and does not allow indications about languages of other services. A language attribute would allow a machine to filter or find radiocommunications or broadcasts which are in a language of interest.

### 2.3.12.3. categoryOfMaritimeBroadcast

Some additional enumerations could be useful.

- Wave height forecast - used in Canada and likely other areas where waves can be a problem (not a tsunami)
- Notices to Fish Harvesters
- ODAS Weather Buoys?
- Ice hazard bulletin
- Ice boundary information
- Ice forecasts (there is 3 - ice report - but that might not be the same; there is 18 - icing warning - but that is for vessel icing hazard not an ice forecast)
- Iceberg Bulletin
- Water level readings
- Weather watches
- Marine weather statements?
- Technical marine synopsis?
- Extended marine forecasts

### 2.3.12.4. Frequencies and Radio Emission codes

There are frequencies which include coded values related to class of emissions such as 'J3E' or 'F1B'. See [https://en.wikipedia.org/wiki/Types\\_of\\_radio\\_emissions](https://en.wikipedia.org/wiki/Types_of_radio_emissions)

There does not seem to be an attribute to encode this information in the S-123 radio communications. **frequencyShoreStationReceivers** and **frequencyShoreStationTransmits** are integer types. Perhaps a new attribute is needed to encode the emission code.

### 2.3.12.5. Correlation of channel numbers and frequency pairs

There is a table in RAMN showing frequencies with emission codes.

The channel numbers correspond to specific frequency pairs. This can be achieved manually by constructing the same number of channel numbers as frequency pairs, leaving some channel numbers blank. However, it has to be intentionally achieved and systems may not expect or interpret the correlation as intended. It would be better to add channel number to the frequency pair complex to allow this encoding to be more explicit.

### 2.3.12.6. Channel as text or number

In the current FC **communicationsChannel** is a text type. Recommendation that this should be a number or integer field.

### 2.3.12.7. Encoding of facilities available

Add guidance to DCEG or perhaps category of facilities to indicate information such as: "Ch24, Ch26, 2142 and 2206: Facilities are available for connecting ships directly to the commercial telephone system on shore."

### 2.3.12.8. Date range of Broadcast

One broadcast may be served by multiple Radio Stations and Radio Service Areas. Also note that some of the content is periodic.

There is a **fixedDateRange** for the **RadioServiceArea**. This means that a new copy of **RadioServiceArea** would be needed when there is a **radioCommunications** that has a unique date range. There should be a **fixedDateRange** within **radioCommunications**.

#### 2.3.12.9. Recommendations: radiocommunications related

- Consider splitting **radiocommunications** into communications vs broadcasts.
- Consider using information types to allow sharing of communication information across **RadioServiceArea** and other features such as **WeatherForecastWarningArea**.
- Consider enhancing radiocommunications to better handle different service languages on different frequencies.
- Enhance **categoryOfMaritimeBroadcast** to model more types of broadcasts.
- Consider new attribute for emission code.
- Consider adding channel number as optional attribute within **frequencyPair**.
- Consider changing **communicationsChannel** to a number field to avoid spurious characters.
- Consider guidance or attributes to indicate facilities supported on specific channels .
- Consider date range to **radiocommunications** for cases where some portion, channels, content is only available during a period.

#### 2.3.13. Contact Details

##### 2.3.13.1. Delivery points

There are multiple **deliveryPoint** entries needed to encode the full contact details but no mechanism/property to distinguish them from one another. The order can be defined but there is no machine readable way to isolate a contact label from a post office box or street address.

##### 2.3.13.2. Language

It is useful and necessary to be able to offer information such as contact details and instructions in different languages.

Add a language attribute to the top level and clone the Contact Details for each language but with the appropriate language in all the sub-textual fields.

##### 2.3.13.3. Recommendations for Contact Details

- Add an optional label or category to delivery point
- Add language indicator to **ContactDetails** to allow managing and finding the details in the language of choice.

#### 2.3.14. Service Hours

Entries in the publication have text such as 'Hours: H24'. The entire **ServiceHours** information type seems pretty daunting to handle when all you need is to say is 24-hour service. Perhaps there could be a Boolean for 24-hour service to make it easier to work with. Maybe the Boolean could be carried on features like Coast Guard Station, Radio Station and Radio Service Area so as not to require the creation/link to a service hours information type.

##### 2.3.14.1. Recommendation:

- Consider simplifying service hours to support common situations such as 24-hour service.
- Consider if **scheduleByDoW** needs to be mandatory

#### 2.3.15. Periodic Date Range, Fixed Date Range

##### 2.3.15.1. Attribute order

This is a complex attribute with sub-attributes for date start and date end. However, in the current Feature Catalogue the order of the sub-attributes is **dateEnd**, **dateStart**. The default order in which the attributes are listed in the GUI for data view/entry and in the GML is the order which is defined in the FC. It is confusing for users because the natural flow would be from start to end. It will cause data entry errors where the end date gets entered as the start date and vice versa. Additional validations could be useful but a simple improvement would be to change the order of the sub-attributes to be more in line with the logical order of the data.

##### 2.3.15.2. Fuzzy dates

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

Consider an attribute within the **dateRange** complex to indicate an approximate range.

### 2.3.15.3. Seasons can vary

Although there are usually calendar dates for when Spring or Summer starts it is more about the conditions which can vary from year to year based on when snow/ice forms or melts.

In this case it is more about the information being applicable to current conditions, Ice Reports if there is any ice. Consider allowing a season name rather than an explicit date range. Winter, Spring, Summer, August.

Also, there are sometimes warnings or information that is relative to other 'seasons' such as 'Lobster fishing season' or 'tsunami' season.

### 2.3.15.4. Recommendation – related to date ranges

- Reorder the sub-attributes of **periodicDateRange**, **fixedDateRange** and **surveyDateRange** to be **dateStart**, **dateEnd**.
- Add an attribute to date range to indicate the range is approximate.
- Consider an optional attribute for a 'season' name rather than a date range.

## 3. GML Data Format

The GML encoding is generally consistent but from an implementation point of view and considering machine readability there are issues to consider.

There is a paper and sub-group under the S-100 WG reviewing S-100 Part 10b GML encoding in order to simplify the encoding patterns and reduce costs, complexity and validation issues that will be faced under the current specification. The recommendations for improvements are intended to improve consistency, simplify implementations (producing and using) and improve interoperability within S-100 and with other non S-100 environments. For example there are several options in the encoding such as having two ways to encode associations. The net result is the same in terms of exchanging content but it means systems need to handle either method. It would be better to select one pattern to use rather than having to develop software and testing for all the permutations and combinations that could arise.

As the work of S-100WG finalize, there will likely be a need to review and update the GML schema used by S-123.

## 3.1. Dataset structure

In S-123 the Dataset structure definition indicates the following in human terms:

The top level Dataset element will contain the following in the following order:

1. gml:boundedBy (coming from gml:AbstractFeatureType)
2. DatasetIdentificationInformation (optional)
3. DatasetStructureInformation (optional)
4. S100:Geometry (optional or repeating) to list all the points, curves and surfaces that can be used by one or more features. This is optional which would indicate that geometries could be shared by reference or carried inline on the Features. In keeping with 8211, the shared geometry should come after the information types.
5. Choice of imember and member (optional or repeating). This indicates that imembers and members could be listed in any order within this section.

### 3.1.1. Recommendation – metadata container

Use a 'metadata' container element to hold DatasetIdentificationInformation and DatasetStructureInformation elements. This probably should be done at S-100 level and would necessitate a change proposal for S-100 Part 10b.

### 3.1.2. Recommendation – reorder elements

It is recommended that a single optional imember collection element be used before the Geometry to carry all the information elements.

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

This would make it more consistent with the S-100 8211 encoding where the information types are in the export before the geometries and features that may be referencing them.

A single member collection element could be used in place of section 4 to carry all the feature types.

### 3.1.3. Recommendation – reduce or remove use of imember/member elements

The proposal would be to change these to be collections instead of single element containers or just get rid of them altogether and just include the information types and feature types at the top level within the Dataset.

## 3.2. Other GML considerations:

### 3.2.1. Choice of enumeration values

Enumerations could be written as strings or as a numeric code representing the enumeration entry.

Numeric:

- The numeric codes take less space but if space is the goal then a binary file is likely smaller.
- Numeric is easier for conditional portrayal if value = N but N could be a number or string and the logic would still work strings just take more space.
- Need to look up string value if value needs to be shown or used in output.

String:

- The string values make the gml more human readable and may simplify XSLT or other conversions.
- String values give consistent results to both enumerations and coded lists.
- Need to use a specified string in a fixed language such as in the official FC.
- User may want to see values in a preferred language. Enumeration codes may be easier to map to strings in different languages.

#### 3.2.1.1. Recommendation – specify enumeration encoding

There does not seem to be a machine readable mechanism for a producing/using the system to know what to expect for enumerations. It would be preferred to either provide one way or define how a system determines what the choice is for a given product. This should be dealt with at the S-100 level and would necessitate a change proposal.

### 3.2.2. Association encoding

Associations can be defined generically such as:

```
<s100:informationAssociation gml:id="ia2" xlink:href="#cn1" xlink:arcrole="http://www.ihonet/S-123/roles/theRegulations"/>
```

```
<s100:featureAssociation gml:id="fa1" xlink:href="#da1" xlink:arcrole="http://example.ihonet/roles/rolea"/>
```

However, the readability and ability to make XSLT templates or filters is more complicated than if the association role is used as the element name such as:

```
<s123:serviceProvider gml:id="f4" xlink:href="#da10" xlink:role="http://www.ihonet/S-123/RadioStation" xlink:title="ServiceProvisionArea"/>
```

With the role as the element name, a simple element XPath or template can be constructed to handle different associations. Also, including the type of the referenced object in the role and the name of the association in the title, provides more information to make decisions without having to traverse the link to see what is on the other end. In S-100, the same role can be used for different association classes and two features could have multiple associations for different purposes so having the information in the link allows this distinction to be made when parsing/using the data.

#### 3.2.2.1. Recommendation:

Use association roles as the element names. Provide role and title for feature and information associations, not required for geometry links.

### 3.2.3. Geometry shared or inline

Note: FOR REASONS OF ECONOMY, DELEGATES ARE KINDLY REQUESTED TO BRING THEIR OWN COPIES OF THE DOCUMENTS TO THE MEETING

Clarify shared geometries not inline with actual examples. Currently the specification is written to specify shared geometry but the xsd allows both shared and inline and the sample uses inline.

#### **3.2.3.1. Geometry encoding**

The elements for geometry and S100:pointProperty appear to be redundant.

#### **3.2.3.2. Recommendation:**

Clarify if geometry is to be shared or inline and if not fixed, how to determine which to use when.

## **4. Exchange set metadata**

### **4.1. Schema on Edition 3**

The main issue with this is that the software was updated to make Ed 4. Some parts are customizable in Ed 4, such as extra fields, and can be ignored but it may be necessary to manually edit the sample file if the Ed 3 metadata cannot be handled as expected. It is expected that a new S-123 will be made in compliance with S-100 Edition 5 where metadata is being changed again and so it is expected that compliance with S-100 Edition 3 may not be a critical thing for production.

For consistency of software support (production and use systems) it will be better to harmonize S-123 with S-100 Edition 5 when available.

### **4.2. Dataset Discovery**

#### **4.2.1. signature**

Note: digitalSignature is optional but probably should not be for official datasets.

#### **4.2.2. Specific usage**

The property **specificUsage** is mandatory from S-100 metadata, but for a product such as S-123 the product is typically scaleless and applies to every usage where someone wants to use it. S-100 only has it as a character string, S-101 limits to three choices (Port Entry, Transit, Overview) Recommendation is that it should be optional or there should be a choice for 'any' when there is a one size fits all product. S-123 states "brief description of the resource and/or resource series usage" but it would be better to make this a fixed value or a choice. This should be dealt with at the S-100 level.

#### **4.2.3. Vertical datums**

Vertical datum and Sounding datum are optional in S-100 Ed 4 but in S-123 they appear as mandatory (because they were in S-100 Edition 3) with a comment that they should be set to Nil as they are not applicable to S-123.

#### **Recommendation:**

It would be better to leave vertical datum and sounding datum out of S-123 than to define them as Nil. Updating S-123 to S-100 Edition 4 or 5 should fix this.

#### **4.2.4. Agency**

Agency information needs to be filled out in multiple places in the metadata. This information should be rationalized in some way to make it simpler. Perhaps the agency in the dataset metadata and contact details could be optional and only needed if it is different than in the exchange catalogue metadata.

#### **4.2.5. Sample Metadata Issue 1 – S-122 instead of S-123**

S-123AppD-1\_EN\_GMLFormat\_Ed1.0.0\S-123 App\_D-1 GMLFormat  
1.0.0\S100\_3\_0\_0\S100Catalog\20170831\S122ExchangeCatalogue.xsd

Assume this should be S123ExchangeCatalogue.xsd. Renamed all instances of S122 within the file to S123.

**Conclusions**

The exercise of creating national S-123 data coverage has been very insightful for identifying gaps and corrections to be made in the S-123 Product Specification and other documentation. It has also highlighted a number of S-100 issues that should be corrected to make future product specification development more streamlined.

**Action Required of NIPWG**

The NIPWG is invited to:

- a. note this paper and take any action appropriate.