Paper for Consideration by S-100WG

Optimisation of the conversion of S-57 data to S-101

Submitted by:	IIC Technologies
Executive Summary:	Informative paper describing current ongoing work in the conversion from S-57 ENC data to S-101 form
Related Documents:	IHO S-101 Product Specification
	IHO S-101 DCEG, S-57 UOC
	KHOA Paper S-100WG Singapore

Introduction / Background

Following the S-100WG meeting in Busan in 2018 a piece of work co-sponsored by NOAA and KHOA was commissioned to look at the process of conversion of S-57 ENC data into S-101 form.

Against a background of the ongoing development of the S-57 to S-101 converter the following objectives were identified:

1. An examination of the current converter application, its operation and results gained.
2. To systematically look at how S-57 could be "optimised" to prepare data for conversion to S-101 and what the scope for performing this optimisation is in advance of S-101 migration
3. Report, summarise and suggest next steps for the conversion process.

An initial report was presented at the S-100WG TSM meeting in Busan and reviews and refinements made to the report following that meeting. A number of subsequent objectives were identified from the initial report and progress towards those have been made. A broader review is now encouraged, after which a final version of the report, together with accompanying documentation and data will be made available to the S-100WG community.

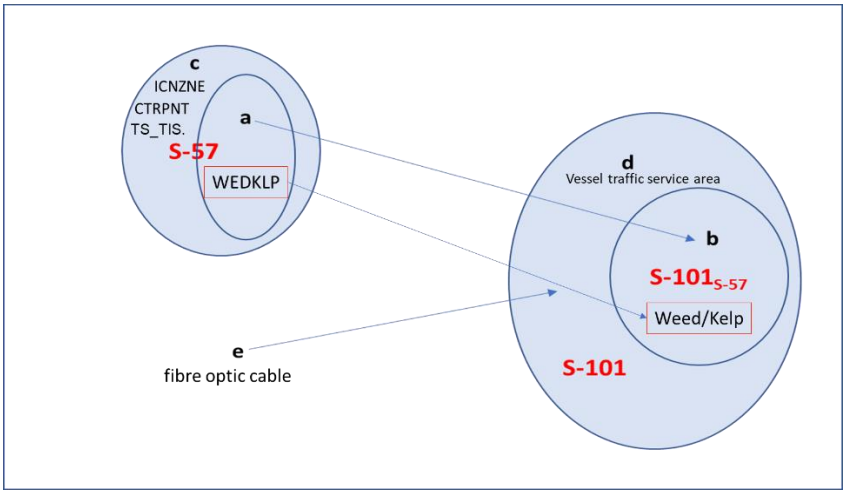
Analysis/Discussion

A systematic methodology was identified and defined during the process of the investigation using up to date toolsets. These were:

1. The ESRI converter tool at its major release
2. The SPAWAR S-101 viewer which also contains numerous analysis tools for data inspection.
3. Data to support the study was made available by several member states. Support from S-100WG members was also gratefully received during the work and the writing of the subsequent report

The core of the study's activities was a comparison of the S-57 Use of the Object Catalogue (UOC) with the current version of the S-101 DCEG. These documents contain extensive and detailed descriptions of the encoding for "equivalent" real world features relevant to the purpose of ENC data arranged in semantic groupings. The study of the two encoding guides was further supplemented by interpretations of the UOC by participating member states – this common practice details how individual member states adapt, enhance (and occasionally) restrict the encoding of ENCs depending on their own internal policies, regulation, cartographic practices and other constraints.

The methodology examined all instances where a “real world feature” was equivalently encoded in the two documents and what, if any, transformation process was required in order to produce S-101 versions of the existing S-57 data. In many cases the S-57 data is directly transformed into an equivalent S-101 feature (and attribute). There are, however a number of gaps, best exemplified in the following diagram:



In the above diagram the domain of features defined by IHO S-57 UOC is represented in the left hand set. The right hand set shows the domain of features covered by IHO S-101 (these can be any of the components of the product specification – features, attributes, associations, feature bindings, enumeration values etc etc…). The elements of the conversion between S-57 and S-101 are categorised as follows:

a. Features which translate directly from S-57 to S-101
b. S-101 equivalents of S-57 features
c. Features which can’t be translated into S-101, i.e. that have no equivalent. The examples given are Control Points and Incineration Zones
d. Features for which related features may have existed in S-57 but which are modelled in a fundamentally different way or enhanced within S-101 – the example is Vessel Traffic Service areas. There are a number of these aggregated groupings within S-101
e. Real world features which were never modelled in S-57 but which have S-101 representations – these are items which were “missed” for a number of reasons from the S-57 view of the world. The example given is fibre optic cable.

A notation for features in both S-57 and S-101 was developed during the writing of the report which facilitates side by side comparison of data before and after conversion. This was used to examine the many instances during the process of the UOC/DCEG comparison. :

$S57f_1$	$S101F_1$

CANALS:
{
OBJNAM = Snapper Creek Canal
SCAMIN = 259999
}
	Canal:
	{
	featureName:
	{
	displayName=0
	language=eng
	name=Snapper Creek Canal
	}
	scaleMinimum=259999
	}

This systematic representation of the S-57 universe and the S-101 universe yields a number of interesting observations which are the output of the report. An executive summary is presented here as initial conclusions. Currently the report is awaiting a broader review from S-100 and S-101 stakeholders and will be finalised in early 2019 following receipt of any further review comments and completion of a number of follow on tasks:

Conclusions

1. The current toolset works! It is possible today to convert S-57 data to S-101 and view and inspect it within a viewer. This in itself is a major achievement and enables the more detailed consideration of the underlying data to take place. The current converter has been enhanced to the point where detailed reporting of its results are output and optimisation of data for the conversion is possible.
2. A converter forms part of an interlocking set of tools and standards in the ENC world which generate, validate and format ENC data. A converter will form part of an interlocking set of tools for supporting the transition from S-57 to S-101 and, properly configured, should form the basis for production of carriage compliant ENC data in both formats.
3. The main source of ambiguity in current S-57 encodings is the use of the INFORM attribute. INFORM has been used for many years to encode elements of real world features (often as a part of member state specific encoding guidelines) which don't "fit" exactly with UOC guidelines or, sometimes, within the UOC where S-57's feature catalogue don't have the capability of representing data adequately. A sample of approximately 900 cells showed approx. 66,000 INFORM encodings
4. Conversely, INFORM can be used to populate many new S-101 features by adopting a structured encoding format and ensuring a suitably configured converter is available to perform the appropriate transformation. The current converter is not designed with such mechanisms in mind but does convert data to a minimum standard of compliance with the S-101 product specification and, to a large degree, the S-101 feature catalogue.
5. There are a number of extraneous features encoded in current S-57 cells which are there to ensure display of data in ECDIS Base mode. As such these are not required in S-101 and could be removed.
6. Using a combination of structured INFORM encodings a suitably configured converter would be able to convert S-57 data into an "equivalently safe" version for use within an S-101 ECDIS. These terms would need to be made more precise (and suggestions for doing that are in the recommendation section of this paper.

Recommendations

1. No abstract definition exists of the actual operation of the current converter. This restricts the ability of any study to be systematic. This would not be difficult to do and is one of the current follow on actions under way following the initial study conclusions. Given the complexities of the DCEG encoding this definition needs to be able to properly define the conversion of features to attributes, C_AGGR to associations and the definition of new features from structured attributes. Additionally, elements like Skin of the Earth feature substitution may only be possible to describe in the abstract due to the complex geometric processes involved in their conversion.
2. No equivalent level of e.g. S-58 compliance exists for S-101 data yet. This restricts the ability of a data producer to evaluate whether a converted cell is "safe" (in the navigational sense) or not. This will follow in time as the S-101 ecosystem matures.

3. The S-101 feature catalogue restricts heavily feature attribute bindings which are commonplace in the existing global ENC dataset. The quantity of these should not be underestimated and, although arrived at through a very systematic and thorough process within the S-101 project team the impact on ENC data remains to be seen. The absence of a feature catalogue for S-57 data makes it difficult to automate the comparison and quantify the impact on individual cells. A version of the S-57 feature catalogue is also a future development proposed following the initial report. This would allow a fully machine readable and automatable statement of S-57 data, its transformation into S-101 and the S-101 data resulting from such a transformation.
4. A graduated level of sophistication of operation of the converter could capture a number of scenarios for S-57 to S-101 conversion in a progressive way. The process of conversion should ideally be considered alongside various different models of how member states intend to migrate from S-57 to S-101 and the assurances required that data remains safe for primary navigation., These levels of compliance resemble “Technical Readiness Levels” proposed elsewhere in the S-100 ecosystem.
5. Preparation of a comprehensive test dataset would be a necessary first step to test that different converters behave equivalently and that the specification for conversion defined is followed correctly. An example test dataset has been prepared as a follow up action to this report which can be demonstrated to illustrate the concept.

Action Required of S-100WG

1. Note the report generated by the initial phase of the converter work and the steps forward made in this area.
2. Note the usefulness of machine readable catalogues for specifying S-10x data and the flexibility it allows when looking to convert between different product specifications. This has applications for broader use of marine geospatial data in the MSDI context.
3. Review the existing paper and present relevant comments.