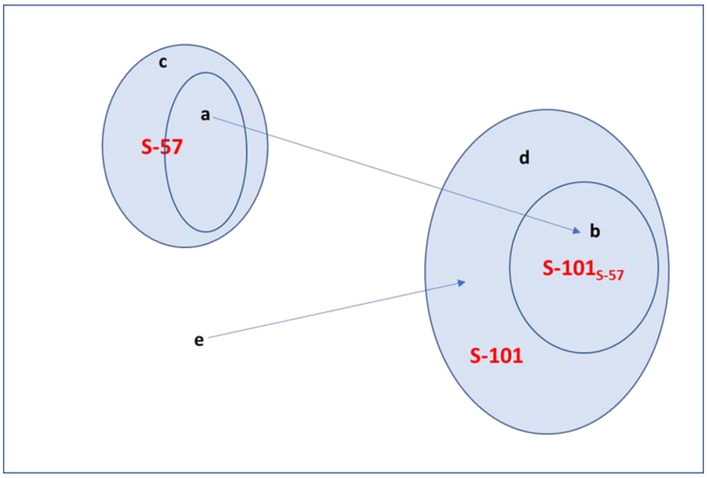
Consulting Report (*Draft*):

Optimised Encoding

For S-57 to S-101 conversion



Submitted to: NOAA

IIC Technologies Ltd
Part of the IIC Technologies Group
The Catalyst
York Science Park
Baird Lane
York

E-mail info@iictechnologies.com
Website www.iictechnologies.com

1. Introduction, background and layout.

This document describes a systematic approach to the subject of S-57 to S-101 conversion. It attempts to look at a high level at the various issues raised by such data transformations and how current S-57 encodings may be modified to "optimise" the S-101 data from conversion.

Background and Previous work.

The process of defining and specifying IHO S-101, the replacement for the IHO S-57 ENC product specification has taken many years to come to fruition. Borne out of a need to correct, enhance and optimize the existing standard, S-101 puts marine ENC data on a much firmer foundation with respect to the overarching geospatial ISO19xxxx standards and supplements the documentation with a comprehensive geospatial registry located at the IHBB in Monaco.

S-101 itself offers many new benefits, both to data encoders (mainly National Hydrographic Offices) and end users. Most end users of ENC are currently ECDIS users, the prime use case for ENC data but an emerging and growing market of end users now exist wishing to use ENC for non-navigational purposes and they are also benefitted from the expanded possibilities of S-101 data mainly due to its flexible nature, its alignment with the ISO geospatial standards framework and the more logical and structured approach to the content encoding.

A number of previous pieces of work were used as a starting point for this report and are available within the S-100 community through the S-100 and IHO website.

The layout of this study.

This report is aimed both at readers wishing to gain an overview and an appreciation of the main issues of S-57 to S-101 conversion. There is an executive summary of the main recommendations of this report contained in "2 Executive summary of outputs and recommendations."

Explanatory detail has certainly not been left out though and the intention is that this report forms a foundation for a practical approach to the many aspects of S-57 to S-101 conversion at a practical, cartographic encoding level. The sections of this report containing more specific information are contained in section 3. Where appropriate recommendations for actual encoding changes are required for optimised conversion these are suitably labelled in the text. There is also a final section detailing some suggestions for further work, some perspectives on the current converter used during the study and topics for papers which may be useful for the working groups charged with further development of the standard and surrounding infrastructure.

Contact

IIC's point of contact for this report for any questions or queries is:

Jonathan Pritchard
IIC Technologies Ltd
The Catalyst
York Science Park
York
YO10 5GA

jonathan.pritchard@iictechnologies.com Tel: +44 7464371695

1.1. Aims and Objectives.

This section explores some of the detail of the objectives of the study and sets out some high level questions to be answered in its execution:

What are we trying to achieve?

A SYSTEMATIC APPROACH – This means taking an approach to the subject of conversion which deals with every conceivable aspect of it so that no elements are lost. As will be seen in the report this overall requirement has necessitated a "top down" approach in defining the technical steps carried out. A systematic approach has led to consideration of the entire S-101 DCEG and feature catalogue and extensive trials of converter technology at a large scale in order to test every eventuality.

OPTIMISED ENCODING – This overall objective is in two parts.

- We attempt to answer the question whether it is possible to "optimise" the encoding of existing ENC data so that when it is converted to S-101 ENC it results in a "better" S-101 ENC.
- If such a process is possible, then detailed guidance on how such optimisation should be carried out, the nature (and scale) of the edits required in the current ENC portfolio, and the effects/benefits of them. As a corollary we also try to answer the question whether there are S-101 elements which cannot be defined/used/optimised by editing the current ENC portfolio.

WHERE ARE WE NOW) - An overall look at the current process and technology of "conversion", how it works, what it does and where any "gaps" may be. Should any changes be required to either the current converter technology or to the surrounding S-101 infrastructure, registry, feature catalogue etc. in order to make the conversion process smoother and more automatic.

A systematic approach.

Finding a systematic way of approaching this problem is not easy. There is no simple, algorithmic way of determining the completeness, precision or functionality defined by an ENC conversion process because of several factors:

An ENC is not a linear "thing" – it doesn't have a start and an end which can then be followed systematically to look at every individual element. Additionally, an ENC is not just defined by its feature content but all its metadata, scale settings, attribution of geometry as well as features and also by the interaction of the resultant ENC database with end user equipment (ECDIS). It is difficult to define whether an ENC encoding is "better" than another – even the concept of an "equivalent" ENC (i.e is an S-101 version of an ENC "equivalent" to its S-57 version?)

The systematic approach followed in the compilation of this report is as follows:

1. Set out at a high level how conversion is carried out and the main factors affecting the outputs from the conversion process. A high level structure is defined showing ENC information, its encapsulation in S-57 data, the conversion and the resultant information encapsulated in S-101.
2. Examine in detail the UOC and DCEG documents side by side and how the two documents approach the encoding of real world features into the language defined by the feature/attribute catalogues of each domain. As we progress through the documents, detail how existing S-57 data can be altered to optimize its conversion to S-101
3. An exploration of how geometry is approached and the equivalence of the two geometric models including attribution, heights, depths and the various edge possible edge cases.
4. Other aspects of the conversion process which may not fit into the original categories defined, including where modifications to the existing regime may be required in order to enable a better conversion process.

Some thoughts on "conversion"

A helpful analogy at this stage may be to think of these things in terms of language translation. Imagine trying to translate a book of instructions or a recipe book from one language into another. When such a book is translated into another language there is a need to make sure that :

a) The "content" is the same – so 500g of flour translates to 4 cups of flour (US).
b) Ensuring that the translation takes advantage of any language-specific advantages that may exist. So a recipe book translation from English into Danish may wish to use

the word “Natmad[1]” or English to German, “Kummerspeck[2]”, words for which no equivalent simple English construction exists.

Whenever a feature is “converted” we can define the concept of the conversion being lossless and an S-57 feature and an S-101 feature being “equivalent” (this is useful because once you can define equivalence you can also define a lack of equivalance more precisely).

Some S-57 features and attributes have equivalents in their respective feature catalogues. These can be seen in the acronyms and abbreviations used for each feature. ${}^{S57}f_1$ is semantically equivalent to ${}^{S101}F_1$ if all the attributes are converted without change. So, in the table below the two features pre and post conversion are equivalent. This is because the intrinsic information content of the two features is the same. Note that we are not saying that the geometry of the two features is equivalent (at this stage) (this is easier to define separately and will be documented later).

${}^{S57}f_1$	${}^{S101}F_1$
CANALS: { OBJNAM = Snapper Creek Canal SORDAT = 200806 }	Canal: { featureName: { displayName=0 language=eng name=Snapper Creek Canal } scaleMinimum=259999 }

You can also, then, define an equivalent conversion even when the feature is converted into two (linked by association) features, as in the case below where a feature has a linked supplementaryInformation feature, e.g:

${}^{S57}f_2$	${}^{S101}F_2$
BOYSPP: { BOYSHP = 1 CATSPM = 27 COLOUR = 1,11 COLPAT = 1 INFORM = Danger shoal OBJNAM = Miami Springs Boat Club Shoal Buoy North STATUS = 8 SCAMIN = 179999 }	BuoySpecialPurposeGeneral: { buoyShape=1 categoryOfSpecialPurposeMark=27 colour=1 colour=11 colourPattern=1 featureName: { displayName=0 language=eng name=Miami Springs Boat Club Shoal Buoy North } status=8 scaleMinimum=179999

[1] Traditional food served at the end of a party, before you’re sent home.
[2] Weight gained through overeating after a failed romance (literally “Grief Bacon”)

	`}` `additionalInformation provides` `{` `    SupplementaryInformation:` `    {` `      language=eng` `      text=Danger shoal` `    }` `}`

1.2. Methodology and Document Layout

The idea of having a methodology is to achieve a good coverage over the study's subject matter and the driving goal of this study is to be "systematic" – this section presents more detail on the study's methodology and its core concepts.

As discussed in the previous section ENC data is fairly difficult to be systematic with because its data is not simply lists of features, data or values. ENC represents an encoding of real world objects where subjective judgement is used to "classify" and assign a number of values to data "objects" (object in S-57, features in S-101 but essentially the same thing).

Additionally relationships between features which reflect real world (and non-real) associations are defined. There is, however, a continuum of complexity within the defining mechanisms and this can be used to broadly define the methodology adopted within the study uses this to drive the activities.

Using Sets to show the conversion process.

In the following diagram the entire feature content of both S-57 and S-101 are shown as abstract sets and the arrows represent conversions between them. In the diagram the S-57 and S-101 contain all a dataset's content itself, i.e. the encoded features, their attributes, geometry, metadata, associations/aggregations and (in the case of S-101, the actual catalogue itself) – the idea of defining the diagram was to focus on the individual elements of the actual datasets themselves as it helps to capture at a high level what the conversion process is doing and trying to achieve:

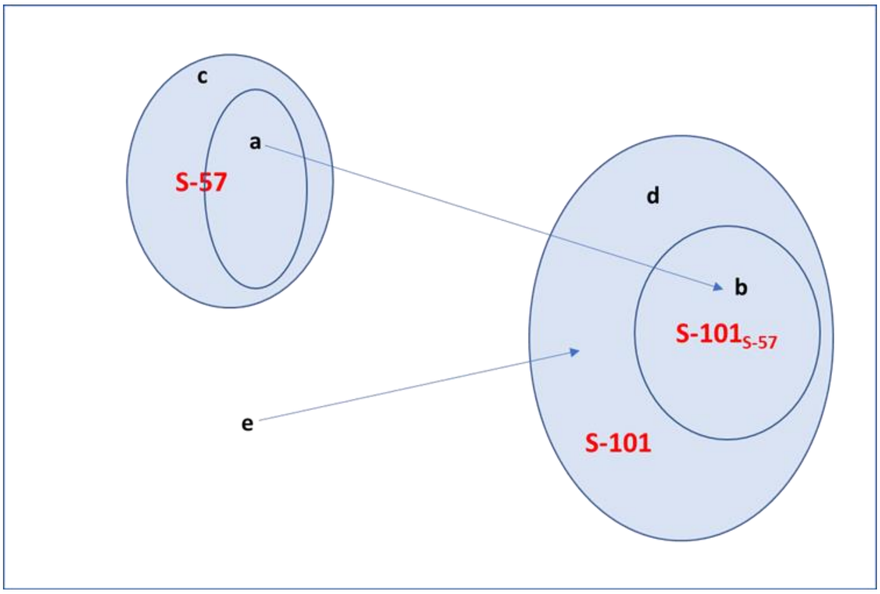


Figure 1: ENC content and conversion in sets

The individual elements of the diagram are as follows:

a. Things in S-57 which can be translated into an S-101 equivalent without loss of information. These can be features, attributes, metadata, geometry – any content of a given S-57 dataset is included in this set including the links they may have with each other (a more detailed notation for these is given in the next section). It can also encompass individual enumerated attributes of particular features (as we will see there is a large quantity of enumerated attributes which exist in current S-57 data but which don't necessarily have S-101 equivalents). Within the feature catalogue these are elements with aliases[3].
b. The S-101 equivalent of (a) – these are S-101 set elements directly translated from the S-57 source
c. Anything in S-57 which can't be (or doesn't need to be) translated into an S-101 equivalent.
d. Anything in an S-101 dataset which has no current defining mechanism in S-57, i.e S-101 which has no feature catalogue analogue in S-57 (as we will see, it may be possible to move content from (d) to (b) depending on both the behaviour of the converter and the encoding of the content (effectively moving it from (d) to (a)).
e. Real world features which previously had no representation in S-57 which are now expressible in S-101 (these are encoded into features (d).

This then suggests a systematic way of progressing the study.

---

[3] E.g LandArea = LNDARE, DepthArea=DEPARE, status=STATUS, direct replacement by acronym.

1. Start with a side-by-side comparison of the features themselves which make up the individual ENCs in both S-57 and S-101 to establish where equivalence can be determined (this is the line from (a) to (b) in the diagram. At this stage simple one-to-one relationships between certain features can be defined (and indeed many features translate directly into their S-101 equivalents). The prime documents for this comparison are the feature catalogues for both S-57 and S-101 and the "encoding" documents (the UOC and DCEG) which define the recommended methods for mapping real world features to encoded data.
2. Look at new elements of the S-101 set defined but with no direct analogue in S-57 (d) and (e) in the diagram.
3. Look in more detail at attribution on elements as defined by the feature catalogue and how feature bindings can be enhanced in S-101 by modification of S-57 attribution. This will also look at "edge cases", so, how S-57 has dealt with real-world phenomena which cannot be readily classified through attribution of features in the catalogue (e) in the diagram.
4. How the two different feature association models differ and the limits of the conversion process in respect of them. These can be prioritised into (a) associations which are required for effective ENC operation or effective portrayal and (b) associations which merely represent real world associations between features.
5. The effective translation of metadata in the conversion process.
6. Other issues – dataset construction, auxiliary files, images, exchange sets etc.

What is "success"?

In order to discuss conversion of ENC from S-57 to S-101 successfully it is necessary to define what "success" actually is. This is by no means a clearly defined goal but the broad aims of such a conversion could be defined as:

1. At a basic level the converted ENC should meet the basic standards of form and content embodied in IHO S-100 and the S-101 product specification. This standardises the format and content of the dataset at an encoding and formatting level and allows for deeper analysis and insights on the outputs of the conversion process.
2. For an ENC to be in compliance with the respective UOC/DCEG pre and post-conversion, i.e an ENC conversion process should not introduce any inconsistencies at odds with the stipulations in the DCEG. If inconsistencies exist in the original dataset (i.e non-compliance with the UOC) then it is accepted that the output of the conversion process may be non-compliant too.
3. The resultant ENC should be compliant with a standardised range of (possibly still to be defined) geospatial and semantic validation tests such as those embodied by IHO S-58. This is also discussed in the "further work" section at the end of this document.
4. A converted dataset should be able to meet the requirements of the IMO Performance Standard for ECDIS. In this context its display (and the ability of the

data to accurately drive display) must meet the standards set out (S-52 for S-57 ECDIS and its equivalent portrayal for S-101) for safe navigation and the content necessary to meet the requirements for Alarms and Indications. Again, this is still under development so a concrete measure of "compliance" is not readily available at this time. The general principle though, is that a "safe" S-57 ENC should convert to a "safe" S-101 ENC for the conversion process to be inherently "safe" itself.

1.3. Assumptions, Notations and conventions.

This study is illuminated by the use of current ENC conversion tools and also by the versions of the S-101 feature catalogue, DCEG and associated S-57 documentation[4]. While the S-57 base is very stable the S-101 documents and tools are understandably in a state of change. All the recommendations within this study for encoding of existing S-57 would need to be re-established prior to a concerted effort to optimise current encodings. As S-101 nears its final publication along with the associated feature catalogue and DCEG the mappings they contain should become stable and usable for the planning of an optimisation programme.

In this document the term "feature" is used loosely to mean any non-geometric entity with an identity within a dataset, so information types, geo features are all referred to as "features". When referring to S-57 the term object is occasionally used. In general upper case acronyms are used when referring to S-57 objects and lowercase (or camel case aliases) when referring to S-101 features to make the distinction precise. Feature/object classes are underlined and coloured blue, e.g LNDARE and Land Area

As part of the systematic approach to this study an independent notation is used to show features, their attributes, metadata and other content of ENC regardless of the ISO8211 encoding in which it is embedded.

This canonical representation of the ENC data is designed to show in a more straightforward way "equivalence" (or whatever has been defined as "equivalence" in any particular context) between S-57 features and S-101 features. An example of the notation of features is shown in the following table:

Representation	Description.
`DepthContour:`	Feature type (as defined in DCEG):
`{ valueOfDepthContour=18.2 scaleMinimum=349999 } {`	Attribute name and values

[4] The section on "Introduction, background and layout. contains a list of references used to create this report.

IIC TECHNOLOGIES

geometry: **{** **id = 120/1812** **orient = forward** **}** **}**	Geometry components by reference (record number within dataset) and orientation.

This notation, though similar to JSON dispenses with some of its features for the sake of clarity The feature name appears on the first line with the attributes listed underneath it. The notation is "pseudo-JSON" with complex attributes appearing on separate lines grouped with pairs of curly brackets. The JSON convention of surrounding keys and values with quotation marks (") has been abandoned, as has the insertion of square brackets around array elements. Associations with either information types or other features can also be shown as separate elements as shown in the following more complex example:

```
Bridge:
{
  categoryOfBridge=1
  reportedDate=20091005
  scaleMinimum=119999
}
{
  geometry:
  {
  }
}

(150/5)
additionalInformation informationProvidedFor
{
  NauticalInformation:
  {
    information:
    {
      language=eng
      text=Railway bridge
    }
  }
}

(100/391)
bridgeAggregation consistsOf
{
  SpanFixed:
  {
    verticalClearanceFixed:
    {
      verticalClearanceValue
    }
  }
```

```
}
```

Where attribute bindings exist in the feature catalogue but no attribute/value pair is defined in the feature instance, the attribute is omitted from the representation. This notation is designed merely to make clear the semantic (i.e non-geometric) properties of features for the purposes of comparison.

Features in general, in their respective encodings are referred to by superscript, e.g

$^{S\text{-}57}\mathbf{F}_1$ refers to a numbered feature (1) encoded in S-57. This makes it easy to show a transformation of a feature from S-57 to its S-101 equivalent, e.g:

$^{S\text{-}57}\mathbf{F}_2$ **->** $^{S\text{-}101}\mathbf{F}_2$

as required. These conventions mean we can consider the transformation of cells to S-101 datasets without being tied up in encoding details and show in a simple and concise manner the content of individual and groups of features.[5]

[5] Currently no encoding-independent machine readable conversion specification exists. One of the recommendations of this report is that a way of encoding the transformation of S-57 to S-101 is established by the IHO working group. Some suggestions which may help in its establishment are included in section “4 Further Work”

2. Executive summary of outputs and recommendations.

Introduction

To avoid for the need for all readers of this report to plough through the fine detail of its outputs, this section is devoted to documenting at a high level the broad conclusions made with respect to ENC conversion from S-57 to S-101 and its current effectiveness against the success criteria defined in Section 1. It also contains a summary of the key output, namely how S-57 encoding may be enhanced to optimise the conversion of the S-57 dataset to its S-101 counterpart. It is designed to be practical in its approach, and to provide actual encoding advice which can be used to optimise encoding of ENCs. Where changes would be required to existing converter technologies in order to enable changes these have been documented as well.

General comments and summary.

Generally the technology available at the time of writing converts S-57 ENCs to correctly formatted and attributed S-101 features encapsulated in compliant S-101 ISO8211 data. Conversion by substitution of S-57 acronyms to their S-101 equivalents and the transformations of their features is largely successful and emerging S-101 cells meet the majority of the tests for success documented. The main area of gaps is the ability to populate the more complicated combinations of new features and attributes defined in S-101. Some of these can be populated automatically such as the new light types, foul ground, wind turbines but some are more difficult to populate automatically, such as bridge spans and range systems.

Additionally, structural changes to attribute catalogues and bindings mean there are large gaps in attribution between existing S-57 database instances and the resultant converted S-101 data (these are all explored in detail in Section 3). There are a large number of existing feature-attribute-value bindings which are prohibited under the current S-101 feature catalogue. While large in number these are the process of a long and thorough period of review by the IHO working groups and represent a best view of inclusion based on navigational significance. They should not be underestimated though and member states may wish to examine in more detail the implication of their removal. Additionally, it highlights the lack of an S-57 version of the feature catalogue detailing features and attributes and their bindings which would easily quantify such differences..

So, while there is a complex process of transformation and a gap in the information contained in the S-57 and its S-101 equivalence the criteria for success outlined (safety, compliance with standards) are largely met with the current toolset (a result of the large effort focused on the feature catalogue and associated DCEG structure). Within this document the focus on optimising existing encoding is population of new features/attributes and those which are clearly articulated within the DCEG.

It is clear that if the current S-101 data is to be enhanced through changes to the existing S-57 then a more complex converter would need to be built. The current converter used in the study converts S-57 to S-101 via substitution of the acronyms together with a set of rules (e.g intelligent aggregation based on C_ASSO, amalgamation of separate S-57 LIGHTS features etc). The general observation on the conversion process made here is that the richer the functionality of the conversion utility and the broader the conversion rules it implements, the wider the set of S-101 data can be produced. So, a section of this report is focused on the nature of those requirements and their broad requirements.

New attributes (some of which may be mandatory (and hence must be included for a compliant S-101 dataset) or have safety/ECDIS display impact are the main focus for encoding "optimisation" in this document. These attributes are identified in detail in later sections. Currently, much of the data defining values of new attribution is included within INFORM attribution of existing S-57 features, either at the suggestion of the S-57 UOC or due to guidance from the ENC producer's local encoding guide. It is suggested that a significant step forward in enhancing the current converter should be to take advantage of the free text nature of INFORM and use it to populate important new attributes. Consideration of metadata features is secondary in terms of conversion. Features like M_COVR, M_QUAL and M_SREL and their S-101 representations can differ substantially in content (e.g M_QUAL and Quality of Bathymetric data) – in most of these cases transformation from other features/attributes is not possible and external input to the conversion process may be needed on a cell by cell basis. Most metadata features only exist in small numbers per dataset, however, and there is little to optimize and little risk of a substantial, manual process of optimising the conversion. A suggested enhancement to the conversion process has been documented in section 3.1 for enhancing the ability to specify metadata features by overriding cell defaults.

Generally, few issues of geometry exist in the conversion from S-57 to S-101. The similarity of the geometry model and its implementation within S-101 are such that the transformation is almost 1-1 (i.e would work both ways (S-101 to S-57 as well) and, aside from the Skin of the Earth changes (discussed in general later in this document) S-57 encoding optimisation is almost all contained within the feature definition and attribution prior to conversion.

The other issue worth noting in general is that the current roadmap for S-101 is not specific about how the migration from S-57 to S-101 is to be achieved globally and this generates uncertainty on the requirements for any S-57 to S-101 converter. Various scenarios are possible for the future rollout of S-101. This disparity of possible future models of S-101 rollout and where S-101 conversion is likely to take place (and how frequently it would take place) has various aspects which require consideration, e.g:

1. Whether the conversion process is a one-time process (so that data can be ingested back into production systems for ongoing maintenance) or on a periodic basis (at new edition or issuing of new updates as part of an ongoing service).
2. Whether feature catalogue attribution bindings should be respected or "allowable exceptions" should be considered
3. How updates are dealt with – is it possible to convert an S-57 update to an S-101 update?

The overall view is optimistic. Although the process of conversion is complex in parts and is the end result of years of work, the methodology is sound, it seems perfectly reasonable that the conversion process can produce "safe" S-101 ENC data (from "safe" S-57 data). Certainly, no "showstoppers" have been identified in the course of writing this report although some areas certainly warrant further work.

The rest of this section concentrates on a summary of some of the main conclusions of the report, suggesting where encoding could be used to optimize conversion of data.

Use of INFORM to create S-101 specific features and attribution.

One of the main recommendations from this report and the majority of the optimisisation which may be done in translation from S-57 to S-101 is to review the content of the current INFORM attribute within existing S-57 data and format it to allow it to be converted into new S-101 features and attribution. The current converter does not support this feature but conceptually this mechanism would allow the resultant S-101 data to be much richer and would enable the broader S-101 feature catalogue to be defined within existing S-57 data.

The core idea is for an enhancement to the conversion process which could take a formatted INFORM attribute and parse it, producing a new attribute value in the resultant S-101 feature.

As well as setting individual attributes on converted features such functionality could be extended to allow new features to be defined from existing INFORM attribution. An example of this could be the creation of CollisionRegulation features from S-57 source CTNARE objects. An example taken from US data is shown in the following table:

S-57	S-101
`CTNARE:` `{` `  INFORM = COLREGS demarcation line, 33 CFR 80.530b` `  SORDAT = 20120103` `  TXTDSC = US5SC34B.TXT` `  SCAMIN = 119999` `}`	`CollisionRegulation:` `{` `  featureName:` `  {` `    displayName=0` `    language=eng` `    name=33 CFR 80.530b` `  }` `  reportedDate=20120103` `  scaleMinimum=119999` `}`

This proposed mechanism of using INFORM to populate new attribution and features has the following advantages:

1. It allows the new, richer attribution of S-101 to be populated from existing S-57 data as required.
2. It uses only the unformatted, existing INFORM attribute within S-57 and therefore requires no bespoke modifications to databases.
3. Much of the information required by new S-101 attributes has been pre-populated by member states already within the INFORM attribute. Common practice for many years has been to populate INFORM with information which does not fit UOC definitions and guidelines and, guided by individual member state encoding guidelines, much of the required data should already exist.

The difficulties of this approach are as follows:

1. It requires modifications to existing converters and the ability to parse structured text within existing INFORM attributes. This would need to be laid out in the converter's documentation and then edits to existing S-57 data would need to be performed to format the INFORM content appropriately. This could, potentially, be a large burden for some member states depending on the choices made in their domestic encoding guidelines.
2. The formatted INFORM attribute would really need to be stripped out of the S-57 data as part of the conversion to avoid confusion in end users. This could potentially lead to a converter which takes a single INFORM-enhanced S-57 ENC and produces a "normal" S-57 ENC (with the encoded INFORM data stripped out) and the new S-101 dataset as part of the same transformation.

Deleted Features.

Several features have been deleted from the S-101 feature catalogue. These deleted features are the result of much work by the DCEG drafting team. A review of the deleted features should be done by any encoding authority to ensure they are comfortable with their deletion and whether they wish to replace the deleted features with other classes.

Creation of new S-101 features

This is dealt with in detail elsewhere but in summary, most of the new S-101 features can be generated from existing S-57 by a suitably configured converter. Classes of creation techniques are:

a. Creation from existing S-57 feature and attribute combinations.
b. Creation from existing S-57 aggregations / associations.

Encoding guidance would be to ensure that aggregation/association features are in place to precipitate the creation and attribution of the new feature. Enhancements to the existing converter would be required in order to ensure this process works correctly.

Dealing with feature catalogue attribute binding changes.

Feature attribute bindings have changed with S-101 after a long and thorough process of review. These could be reviewed in more detail by ENC originators and the underlying justifications for their removal examined prior to any parallel release of S-57 and S-101 data.

Extraneous encoded objects (and in the water attributes)

Display of many point objects is not guaranteed on the ECDIS. For this reason common encoding practice has been to either not encode point objects for some classes, to encode very small Area or Line features or to encode coincident point feature objects which display in ECDIS base such as LNDARE or LNDMRK. Within S-101 attribution to force display in ECDIS base display has been defined (most notably using the inTheWater attribute). This then poses the problem of removal of the extraneous coincident features. Have these point displays been fixed/changed? This needs to be thought out. In some cases it's fine to leave the coincident features in the converted dataset but where S-101 has changed the encoding it is worth removing them.

In order to set intheWater automatically a converter could determine the underlying group 1 feature and set intheWater if it is DEPARE or DRGARE. This would force display of the relevant object in ECDIS Base display when in or over navigable water as required. However, this would require enhancements to existing converters and also an acceptance of the algorithmic setting of the attribute so it may be worth trying to set it from INFORM attribution as well and allowing a converter to set it by configuration parameters. The intersection with underlying group1 features may be more complex if the primitive is area in the S-57 and this should be the subject of review prior to conversion. The features with inTheWater in their S-101 bindings (and in scope for this procedure) are:

1. Building(BUISGL) :point :surface
2. Landmark(LNDMRK) :point :curve :surface
3. Silo/tank(SILTNK) :point :surface
4. Wind turbine( ) :point
5. Fortified structure(FORSTC) :point :curve :surface
6. Crane(CRANES) :point :curve :surface

Two examples of features with coincident encodings recommended in the UOC are listed below:

Named Sea Area / SEAARE: The UOC suggests encoding a Sea Area to coincide with Anchorages containing the name of anchorages because ACHARE-OBJNAM attributes don't display on ECDIS.

Offshore Buildings, Landmarks and Silo/Tank.

Current UOC guidance reproduced in the following Figure recommends the encoding of coincident PIPNT, LNDARE, PONTON features to ensure display when encoding an offshore building, landmark or silo/tank. These extraneous features are replaced in S-101 by the setting of the in the water attribute and so are not needed once the conversion process is complete.

A suitably configured converter could perform a deletion of the extraneous features automatically (using the coincident nature of the point features to determine which ones to delete).

Boatshed. If the service being provided by the structure is known, object classes **SMCFAC** (see clause 4.6.5) or **HRBFAC** (see clause 4.6.1) may also be encoded.

- If it is required to encode an offshore building, landmark or silo/tank, an ECDIS Base Display object (e.g. **PILPNT**, **LNDARE**, **PONTON**) must also be encoded coincident to ensure the feature is always displayed on the ECDIS. Where fitted, lights should be encoded as described in clause 12.8, with the **BUISGL**, **LNDMRK** or **SILTNK** being used as the structure object for the **LIGHTS** equipment object(s) (see clause 12.1.1).
- For encoding offshore windmotors, see clause 11.7.4.

Figure 2: Current UOC guidance on encoding of offshore buildings, landmarks and silo/tanks

Facility (see clause X.X) may also be encoded.

- For buildings located in or over navigable water, the Boolean attribute **in the water** must be set to *True* to indicate that the feature is to be included in the ECDIS Base Display. Where such structures are located over the water it is not required to encode any supporting structures (for example piles, stilts).
- The complex attribute **vertical clearance fixed** must not be populated, unless the building is located over

Figure 3: DCEG guidance on encoding of offshore buildings.

Alternative encodings.

As discussed in the previous section some S-57 features have no display within the current ECDIS S-52 portrayal. According to the UOC these may be encoded as other objects in S-57. Encoders should review their own local guidance and determine if, indeed, alternative encodings were used. If so, they may wish to review whether the ECDIS display is now enabled (e.g through inTheWater attribution) and re-encode these features. This is highly dependent on the encoder's own local guidance however.

For instance US guidance is to encode point RAPIDS features as LNDMRK,INFORM=rapids so there is an opportunity to re-encode these as rapids automatically via the INFORM mechanism outlined earlier in this section.

Other examples where the UOC currently suggests alternative encodings for the purpose of ECDIS display are:

a. PRDARE - Production and Storage Areas (Point) with CATPRA = null).
b. Offshore FORSTC / fortified structures
c. SILTNK
d. TUNNEL (Point – US encode INFORM=tunnel)
e. WATFAL (Point)
f. RUNWAY (Point)
g. GRDIRN / Gridirons - these are sometimes encoded as OBSTRNs so it would be important to know if the OBSTRN nature of them needs to be preserved or not when converting. Local encoding guidance will recommend (UOC specifies OBSTRN/SLCONS or area GRDIRN).

Geometry and Skin of the Earth

There are changes to group 1 (Skin of the Earth) features in S-101. A more detailed look at the implications for dataset conversion is included in Section 3 of this report. The change will necessitate geometric transformations to maintain the planarity of the skin of the earth features but there should be few immediate encoding implications.

The main decision to make is what to replace the deleted group 1 S-57 features (HULKES, FLODOC, PONTON) with (they will need to have underlying skin of the earth features in S-101. Simply replacing them with Unsurveyed area features is a possibility but a "better" encoding may be to (automatically) take on the existing surrounding skin of the earth feature (or features if there are more than one). This would require review of the S-57 objects listed above and would require a converter capable of making such substitutions intelligently.

Geometry conversion.

S-57 to S-101 conversion is mainly concerned with the semantic conversion of the features, their attributes and associations, rather than with the geometric components. If we visualise an S-57 cell as a set of records of individual types, feature and geometric as in the following diagram:

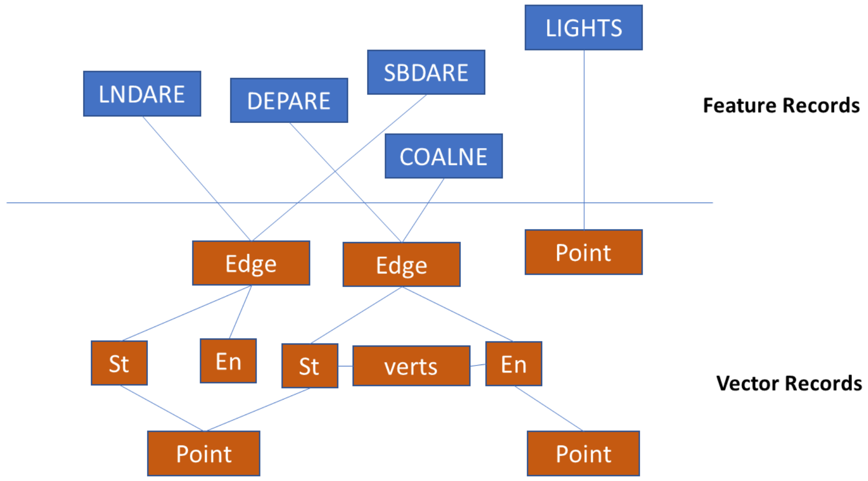


Figure 4: The split between geometry and objects in an S-57 cell

Then it is possible to see the clear separation between the geometric components and the feature components (in this context both meta features and associations (in the S-57 context) are also "features"). This allows for coincident geometry and is the fundamental design principle of S-57's powerful geometry model.

This design has been used for S-101 as well but extended to be more comprehensive, flexible and robust. As the geometry conversion to S-101 only introduces extra layers of abstraction[6]

[6] The introduction of formal composite curves and surfaces allows a more compact model for referencing geometry, particularly when complex/coincident curve and polygon features with many holes are defined in ENC data.

to its model (except for skin of the earth changes dealt with in the previous section) the only changes that need to be considered are the links from the features to their geometry components (the FSPT records under S-57 and the SPAS records in S-101). The only changes to actual coordinates (whether points or intermediate vertices) generated by the S-101 conversion are those generated by the changes to the skin of the earth features (as the planar nature of these feature classes needs to be preserved) so no encoding changes are required to construct a fully functional S-101 geometry model from an S-57 one.

In summary, when considering updates to S-57 in order to optimise S-101 transformation only changes relating to skin of the earth modifications are even relevant (examined in detail elsewhere in this document) and the core focus of optimisation are, hence, to be found in the semantic mappings of features to their S-101 equivalents.

2.1. Recommendations for further work

The following are suggestions for further work based on the outputs from this study. These are explored in more detail in Section 4.

Conversion Specification

Creation of an objective conversion specification which specifies how S-101 features, attributes, encodings and metadata should be constructed from existing S-57 (where possible) and guidance as to how the gaps can be filled.

Distribution and conversion Methodologies

A closer look at distribution and conversion methodologies, where conversion takes place and the implications for converter functionality, distribution processes and end user impacts.

Converter enhancement requirements

Detailed specification of converter enhancements required to carry out the encoding optimisations contained in this report as well as a more systematic investigation of current technology in this area based on structured test datasets.

A Test Area

A worked example within a particular area based on a single member state (US?) and their encoding guides. This would include a draft review of features, “before and after” pictures for particular cells and development of prototype test datasets for investigating converter functionality. This could also include enhanced converter implementation.

Updates.

A detailed look at S-101 ENC updates. Can a converter process updates and produce information on them inserted into the cell? Is it possible to directly convert updates?

3. Detailed outputs.

This section lists the detail found in the course of the study. The outputs from a side-by-side study of the UOC vs the DCEG (with optimised encoding in mind) are listed, with suggestions for how current INFORM attribution can be used to populate new S-101 attribution. There are also sections on how new features and attributes can be populated and a more detailed examination of the geometry conversion aspects of S-101 conversion.

3.1. Detail – Comparison of UOC and DCEG specifications

Introduction.

This section is the result of a comparison between the ongoing DCEG and existing S-57 UOC. The main element in this section is a list of where INFORM attribution in S-57 ENCs could be useful in enhancing S-101 attribution but there also lists of features for which other observations are made and where some encoding optimisation is possible. It should be acknowledged that there are many national encoding guidelines in existence (mostly derived from the UOC) so the amount of optimisation possible by an encoding member state may differ significantly depending on their existing database.

S-57 use of INFORM – Use in transformation.

All the instances listed here show how INFORM could be used to determine new feature creation and/or population of new S-101 attributes. Sometimes current UOC guidance exists to populate INFORM in certain circumstances (e.g range information for local magnetic anomaly data) and sometimes local guidance may result in INFORM already being pre-populated (e.g US guidance is to populate INFORM=rapids for point RAPIDS features (because RAPIDS aren't displayed in ECDIS). The use of the INFORM attribute is not purely for population of attributes but can also define associated feature instances and information types – examples of all of these are included in the lists in this section.

LNDARE/Land Area. INFORM is not mentioned in DCEG, US uses INFORM for translations of names (primarily into Hawaian) These could be used to populate alternative languages in the "Feature Name" attribute.

Population of Nature of surface in coastline features. COALNE (currently the unsurveyed nature of the coastline or the NATSUR details (NATSUR is not currently an attribute of COALNE but Nature of Surface is an attribute of Coastlne)))

RESARE/Restricted Area Navigational - INFORM often contains speed limits which are now encoded into "vessel speed limit".

FSHZNE currently UOC specifies that INFORM should carry value and units of limits of the fishery zone. DCEG guidance is that this should be translated into the Name attribute.

DMPGRD – currently dates are encoded in INFORM

NAVLNE INFORM often carries legend and measured distance information which now have their own attributes in S-101.

HULKES/Hulks – new enumeration values for CATHLK = casino or training vessel could be encoded into INFORM and used to populate CATHLK (GB uses these for training vessels)

CGUSTA, MRCC status is encoded in INFORM and needs to be mapped to "MRCC = true"

ACHARE / Anchorage area – Individual reported anchorages without defined limits are currently encoded as points with INFORM=Reported anchorage. These could be translated into appropriate Anchorage are features in S-101. INFORM / Nautical Information however is also used to store time limits of anchoring.

LAKARE / Lakes – the ability to mark lakes as "intermittent" (STATUS(=5) is not an S-57 attribute of LAKARE) is new in S-101. This could be attributed in INFORM (and probably currently is for some encoders).

Lights – population of major light attribution. Although this could be partially automated (e.g by consideration of the nominal range) the ability to override an algorithmic method and attribute a light as a "major light" would be useful within the INFORM attribute for the individual lights.

Local Magnetic Anomaly / LOCMAG – currently the S-57 attribute VALLMA holds the S-101 equivalent Magnetic anomaly value maximum. There is also an attribute for the minimum value of the anomaly in S-101 and current UOC advice is to encode this in INFORM. These INFORM values could be used to set the minimum (i.e the range of values) as appropriate.

Span Opening. In the encoding for Bridges an association with Service hours and/or Non-Standard Working Day needs to be encoded for opening times. These are currently frequently encoded in INFORM and could be parsed out and linked by association (recognising that this complicates the functionality of the converter.

DISMAR / Distance Marks – current UOC guidance encodes the value of distance within INFORM. Additionally, a frequent value encoded within INFORM is the origin from which the distance is measured. These values map to the S-101 attributes Measured distance value, distance unit of measurement and reference location. Explicit encoding of these values within INFORM would populate these values in the resultant Distance Mark feature. Measured distance value is a mandatory feature. (Examples in `US5FL61M`.)

Berths – INFORM currently contains max permitted draft information (As specified by the UOC), these should be checked in the converter. The Max permitted draft should be populated from the INFORM. DCEG points out a distinction with terminal facilities which should be encoded as harbor facilities which should be checked in the S-57 encoding.

LOGPON / Log ponds – Seasonal Log ponds can now support periodic date ranges. These could be encoded in INFORM.

CBLSUB / Cable Submarine (and CBLARE/Submarine cable area) – many cables have INFORM=fibre optic cable and this could certainly be used to populate CATCBL. All live telecoms cables are now fibre optic (these need to be checked) and disused (including non-fibre optic cables) have CATCBL not encoded (need to check current encoder removes CATCBL from disused features). New attribute values for CATCBL are ferry and fibre optic and ferry cables could be attributed through INFORM in S-57. Existing encoded instances for ferry cables should be checked to make sure they can be translated or encoded. Similarly the contact details for the cable would have to be encoded into INFORM and converted into the associated information type "Contact Details"

Seaplane Landing Area(s) – now also used to encode an area where seaplanes draw water for firefighting. This could be coded into INFORM.

Dock Area / DOCARE – New attribution for the horizontal clearance length and width are present in the S-101 Dock Area. These could be parsed out of the INFORM on the DOCARE source feature.

PIPARE / Pipeline area – contact details are sometimes present in INFORM and could populate the new Contact details attribute (there is no corresponding attribute in S-57). There is an additional constraint in the DCEG over ensuring the outer limits of the area correspond to safe distances which would require validation on the source S-57 prior to conversion.

Features requiring minor inspection.

The following observations on the DCEG/UOC comparison may have a bearing on current S-57 encoding for better S-101 conversion. It is envisaged that these features should be reviewed (possibly automatically) and action taken if necessary. The features in this section may not be able to be parsed automatically although certainly some of the conversion process can be automated. There are also observations on the current conversion process. A more detailed look at selected features takes place in the next section.

LNDARE – topology constraints. LNDARE has an expanded set of topology constraints within the DCEG. These will need checking (although much of this could be automated by a suitable database) and correcting prior to conversion if necessary.

- **Land Area** is usually of type surface; it may, however, be of type point (for example islet, rock that does not cover), or of type curve (for example islet, offshore bar, isthmus).
- **Land Area** of type curve or point must not be encoded on top of **Land Area** of type surface, unless it is also covered by a **Lake**, **River**, **Dock Area**, **Lock Basin** or **Canal** feature of type surface.
- The limits of a **Land Area** of type surface must share the geometry of at least one of the following features:
  - **Coastline**, **Shoreline Construction**, **Gate**, **Dam** of type curve;
  - **Data Coverage**, **Gate**, **Dam**, **River**, **Tunnel**, **Dry Dock**, **Canal**, **Lake**, **Lock Basin**, **Dock Area**, **Land Area** of type surface;
  - **Causeway**, **Shoreline Construction**, **Mooring/Warping Facility**, **Wreck**, **Obstruction**, **Pylon/Bridge Support** of type surface; and having attribute **water level effect** = *1* (partly submerged at high water), *2* (always dry) or *6* (subject to inundation or flooding).

Figure 5: New Land Area topology guidance.

OBSTRN / Obstruction – there are more constraints on obstructions e.g must set HEIGHT on WATLEV=1 or 2, surface OBSTRNs must be encoded as foul area. Platforms are new for capture. A detailed look at existing encodings of obstructions would ensure their conversion is accurate.

DAMCON – a submerged weir is no longer an Obstruction feature so should a Dam, WATLEV=3 be added to alarms/indications map when this is written for S-101?

Built Up Areas:

An extract from the DCEG is shown below:

features of type point.

- Where the source indicates that a built-up area extends into navigable water (over **Depth Area** or **Unsurveyed Area** object(s)), an encoded **Built-Up Area** feature of type area, where required, should be extended over the water area. The seaward edge of the built-up area ("apparent" coastline) must be encoded using a **Coastline** feature having no value populated for the attribute **category of coastline**, and the corresponding spatial edge(s) should have the spatial attribute **quality of position** = *4* (approximate). The actual coastline should be encoded as **coastline**, having no value populated for **category of coastline** and no value for **quality of position** on the related spatial edge(s). Underlying bathymetry (depth contours, soundings) should be encoded as required. Exceptionally, encoders may extend the underlying **Land Area** feature seaward to the "apparent" coastline, which should have the corresponding spatial edge(s) populated with the spatial attribute **quality of position** = *4* (approximate). This encoding should be considered for generalization purposes on smaller scale ENCs.
- For encoding individual buildings over navigable water, see clause X.X.

Built up areas which extend over Depth areas or Unsurveyed Areas should be checked to see that they conform to the new guidance in the DCEG

Conveyors

The DCEG guidance on conveyors is shown below:

Scale minimum	(SCAMIN)	See clause X.X	IN	0,1

INT 1 Reference: D 25

6.8.1 Conveyors (see S-4 – B-382.3)

If it is required to encode a conveyor, it must be done using the feature **Conveyor**.

The value of the vertical clearance between (high) water level and any fixed overhead obstruction must always be given, where known, on the largest maximum display scale ENC data intended for navigation under the

obstruction, and for detailed passage planning. The datum above which clearances are given must be a high water level, preferably Highest Astronomical Tide (HAT), where the tide is appreciable. The value for the vertical clearance must be encoded using the complex attribute **vertical clearance fixed**, and sub-attributes populated relevant to the feature, rounded down to the nearest whole metre (unless under 10m, when metres and decimetres may be quoted). In areas where the tidal range is not appreciable the datum above which clearances are given should be Mean Sea Level (MSL).

Remarks:

- If it is required to encode an overhead cable car, it must be done using a **Conveyor** feature, with attribute **category of conveyor** = *1* (aerial cableway (telepheric)).
- In navigable water, conveyor supports must be encoded, where possible, using a **Pylon/Bridge Support** feature (see clause X.X), with attribute **category of pylon** = *3* (aerial cableway/sky pylon).

Distinction: Cable, overhead; crane; pylon/bridge support.

Figure 6: DCEG guidance on conveyors

This is more detailed than the current guidance within the UOC (little guidance over attribution exists in the current UOC) so careful inspection of these features would be necessary to ensure relevant attribution is defined properly (e.g. datums, resolution on vertical clearance). During review of the conveyors the correct encoding of the supports could be done to ensure they transform correctly to pylon/bridge support features.

Piles / Piling / Area of piles.

Current ENC encoding only allows for Point primitives on piles (PILPNT). DCEG allows for both curve and surface primitives (and introduces two new CATPLE enumeration attributes for these constructions 5=piling (curve) and 6=area of piles (surface). To auto-construct these it would be necessary for a suitably configured encoder to draw outlines or join up individual piles which are aggregated together in the source S-57. This could be identified as a “nice to have” feature for the S-101 conversion but there are numerous examples from the text in current INFORM which shows many instances of multiple piles where a curve or surface feature would be a far simpler representation.

HRBARE / Harbour Area

There is a small note in the DCEG for Harbour areas, reproduced below:

Remarks:

- If it is required to encode a named harbour area over which there is no jurisdictional authority, it must be done using the feature **Sea Area/Named Water Area** (see clause X.X).
- A masked line may be used to suppress the symbolisation of the boundary, where such symbolisation is

Figure 7: Encoding of Harbour Areas

This implies that harbor areas currently encoded in S-57 ENC may possibly need to be re-encoded as Sea Area / Named Water Area if there is no jurisdictional authority. These could be held within INFORM (so that the S-57 feature remains HRBARE) or translated to SEAARE in the S-57. Either way, the encoder should be familiar with which HRBARE’s meet this stipulation and be able to identify them.

Further consideration of selected feature classes

In this section a more detailed look at some selected features is tabulated along with some ideas for their optimisation in existing S-57. This section contains more detailed documentation on certain features and their conversion to S-101.

Anchorages

Feature name	**ACHARE / SEAARE(S-57)**
Remarks	Current UOC guidance is below: and INFORM = Reported anchorage. • For named **ACHARE** of type area, the name (when populated for the attribute OBJNAM) will not display in ECDIS. Where it is considered necessary to display the name of an anchorage area in ECDIS, a **SEAARE** object of type area should be encoded coincident with the **ACHARE** area. OBJNAM for the **SEAARE** should be populated with the name of the anchorage as has been populated for the **ACHARE**.
Current Converter behaviour	
Current	- Not known. Suspect SEAARE is left in post-conversion.
Possible enhancements?	- Convert ACHARE to anchorage, - Remove coincident or intersecting SEAARE (possibly this should be configurable through INFORM actions) - See also later section on how INFORM should be used to capture reported anchorages.
Feature / Attribute Mapping	
Encoding guidance/actions	- Check that names in SEAARE match those in OBJNAM of ACHARE features and that removal of SEAARE causes no issues (e.g multiple names etc) - Check attribution of INFORM for reported anchorages.
Examples:	
Overall Categorisation	Minor.

Marine Farms (MARCUL / Marine Farm)

Feature name	**MARCUL / Marine Farms**
Remarks	Current UOC guidance is below: completely under water. • The attribute **height** must be populated for **Marine Farm/Culture** features having attribute **water level effect** = *1* (partly submerged at high water) or *2* (always dry). • The attribute **vertical length** is used to populate the distance of the marine farm above the seabed Height is now mandatory for Marine farms with WATLEV = 1 or 2.
Current Converter behaviour	
Current	-
Possible enhancements?	-
Feature / Attribute Mapping	
Encoding guidance/actions	- Ensure height is set for MARCUL,WATLEV=1 or MARCUL,WATLEV=2. There are none in current US data but some do exist in other countries' ENCs. These

	would need to be populated to ensure the resultant S-101 dataset meets UOC specifications. - There is an existing S-57 attribute (VERLEN) for recording height so no need to insert into INFORM and parse it out.
Examples:	
Overall Categorisation	Minor.

PONTON / Pontoons

Feature name	**PONTON / Pontoon**
Remarks	be encoded using PONTON. • PONTON objects of type area are part of Group 1. Note that the attributes PEREND, PERSTA, DATEND and DATSTA should not be populated for any Group 1 object classes. A CTNARE object (see clause 6.6) may be used to indicate to the mariner that the presence of a pontoon of type area is temporary or periodic, using the attributes INFORM or TXTDSC (see clause 2.6.1d).
Current Converter behaviour	
Current	-
Possible enhancements?	- - Converter can take PEREND/PERSTA information from coincident CTNARE and populate into PONTON features. Is the resultant CTNARE then needed? Possibly INFORM could be used to control converter behaviour as well.
Feature / Attribute Mapping	
Encoding guidance/actions	- Search for PONTON where coincident CTNARE contain temporal information e.g DATSTA/DATEND and PERSTA/PEREND. - Check that the new boundary conditions imposed by DCEG are respected in the resultant features.
Examples:	
Overall Categorisation	Minor.

ICEARE (Ice Area)

Feature name	ICEARE / Ice Area
Remarks	Current UOC guidance on ICEARE encoding on land and in the sea If it is required to encode the portion of a glacier that is on land, it must be done using an ICEARE object, with attribute CATICE = 5 (glacier) covered by a LNDARE object (i.e. the glacier does not form a hole in the land area). If the seaward edge of an encoded glacier is coincident with the coastline, this edge should be encoded using a COALNE object, with attribute CATCOA = 6 (glacier (seaward end)). ICEARE objects that are located in the sea must be covered by a LNDARE or UNSARE object, if the depth of water beneath them is unknown, or covered by a DEPARE object, if the depth is known.

	DCEG guidance: Remarks: • **Ice Area** features that are located in the sea must be covered by an **Unsurveyed Area** feature, if the depth of water beneath them is unknown, or covered by a **Depth Area** feature, if the depth is known. • As ice fronts move, a date when the limit was surveyed should be included, if possible, using the attribute **reported date**. 5.13.1.1 Glaciers (see S-4 – B-353.8) If it is required to encode the portion of a glacier that is on land, it must be done using an **Ice Area** feature, with attribute **category of ice** = 5 (glacier) covered by a **Land Area** feature (that is, the glacier does not form a UOC allows ICEARE features located in the sea to have LNDARE as the underlying group 1 feature if the depth is unknown. The DCEG does not allow for this and mandates only Unsurveyed area features.
Current Converter behaviour	
Current	- Unknown. Needs test data.
Possible enhancements?	- Converter should replace non-glacier ICEAREs with underlying LNDAREs to Unsurveyed Areas to enforce the DCEG requirement. - Additionally DCEG mandates that quality of horizontal measurement = approximate for visible coastline (COALNE,CATCOA=6) – this extra attribution could be added by the converter when converting the feature.
Feature / Attribute Mapping	
Encoding guidance/actions	- Search underlying group1 features underneath existing ICEARE. Any with LNDARE should be replaced by UNSARE to ensure correct transformation. - Alternatively a suitably configured converter could carry out such substitution. - SORDAT is transformed into “reported date” encoders should be familiar with this and comfortable in this translation.
Examples:	- No examples. Needs test data.
Overall Categorisation	Minor.

3.2. Detailed outputs – Creation of new Features.

Stgrategies for dealing with the population of new attributes is set out later, this section looks in detail at new features and attributes introduced by the S-101 feature catalogue and how they might be formed either from existing data or through INFORM attribution.

This summary details some of the most illustrative examples of how attribution in S-101 can to be populated by a converter. The current converter may produce some of these features (where that has been verified it is noted here). Where conversion is not completely straightforward a longer discussion is included in this section.

Feature	**Notes**

Quality of non-bathymetric data	Populated largely from existing M_ACCY features
Local direction of buoyage	See section following table for detail on this feature.
Quality of Bathymetric Data	Populated largely from existing M_QUAL features
Update information	Won't be populated. Doesn't exist in S-57, would need to be done externally
Island Group	Aggregation of LNDAREs – See following section.
Span fixed	See following section on bridge spans
Span opening	See following section on bridge spans
Wind turbine	See following section on wind turbines.
Depth - no bottom found	Should be populated from SOUNDG,QUASOU=5. These soundings should be grouped together with depth rounded to whole numbers of metres (as per DCEG)
Foul ground	Should be populated from OBSTRN,CATOBS=7
Discoloured water	See following section on discoloured water.
Range System	See following section on Aggregations
Fairway System	See following section on Aggregations
Two-way route	See following section on Aggregations
Deep Water Route	See following section on Aggregations
Traffic separation scheme	See following section on Aggregations
Archipelagic Sea Lane	See following section on Aggregations
Information area	See following section on Aggregations
Pilotage district	See following section on Aggregations
Collision regulations limit	Similar to "Discoloured Water" – see following section.
Restricted Area Navigational	See following section on creation of Restricted Area features
Restricted area regulatory	See following section on creation of Restricted Area features
Light all around	Should be populated unambiguously from LIGHTS where no sectors are encoded and CATLIT != 1 (directional function)
Light sectored	Should be populated unambiguously from LIGHTS features where each S-57 light has SECTR1/SECTR2 defined.
Light fog detector	Should be populated unambuiguously from CATLIT=7
Light air obstruction	Should be populated unambuiguously from CATLIT=6
Buoy Emergency Wreck Marking	Should be populated from NEWOBJ features in S-57. An emergency wreck marking buoy is specified in the UOC.
Physical AIS aid to navigation	Should be included in the INFORM of the aid, e.g INFORM=" Automatic Identification System (AIS) aid to navigation: MMSI,,,,"
Virtual AIS aid to navigation	As per emergency wreck marking buoy this should be populated from NEWOBJ features in S-57. An

	emergency wreck marking buoy is specified in the UOC
Vessel traffic service area	Surface
Text placement	Not populated from S-57

Local Direction of Buoyage

An extract from the S-57 UOC concerning the encoding of local direction of buoyage is reproduced below.

> Within a data set, there may be some areas where the direction of buoyage is defined by local rules and must, therefore, be specified. These areas should be encoded as separate **M_NSYS** area objects, with the attribute ORIENT indicating the direction of buoyage (MARSYS must not be encoded). **M_NSYS** objects with a value encoded for ORIENT must not overlap, but in areas where local buoyage directions apply, **M_NSYS** with a value encoded for ORIENT may overlap **M_NSYS** with a value encoded for MARSYS (see Figure 16 below).

Figure 8: UOC entry referring to local direction of buoyage

As can be seen the new feature of “Local Direction of buoyage” should be encoded whenever an S-57 M_NSYS feature has an ORIENT setting and no defined M_NSYS, i.e

M_NSYS: **{** **MARSYS = 2** **}**	**NavigationalSystemOfMarks:** **{** **marksNavigationalSystemOf=2** **}**
M_NSYS: **{** **MARSYS =** **ORIENT = 248.4** **}**	**LocalDirectionOfBuoyage:** **{** **marksNavigationalSystemOf=1** **orientation=248.4** **}**

So, in terms of the transformation from S-57 to S-101, M_NSYS features can be transformed either into NavigationalSystemOfMarks or LocalDirectionOfBuoyage depending on their attribute values. Encoders should ensure that the MARSYS attribute is set to null and ORIENT set appropriately in order to ensure correct definition of the LocalDirectionOfBuoyage feature.

The DCEG also specifies that the `marksNavigationalSystemOf` attribute should be set to the value for the underlying Navigational System of Marks,

e.g from the DCEG.

Remarks:

- The mandatory attribute **marks navigational – system of** is required for ECDIS portrayal, and must be populated with the same value as populated for the **marks navigational – system of** on the underlying **Navigational System of Marks** feature.

Figure 9: DCEG remarks on Local Direction of Buoyage

So, in order to ensure the features are properly attributed the whole cell should have an M_NSYS set (with an appropriate MARSYS value) and each Local Direction of Buoyage M_NSYS should have MARSYS=null and ORIENT set to the appropriate value.

Aggregation features - Island Groups

Island groups are a good example of where existing S-57 C_ASSO or C_AGGR features can be transformed into the S-101 equivalent aggregation features. There are many S-101 aggregation features which need to be assembled from individual parts and the converter should make an attempt at creating these features from existing component parts if the source S-57 cell has defined a C_ASSO feature between them.

For instance, an Island Group feature would be formed from

The individual aggregation features are as follows:

a) Deep Water Route
b) Fairway System
c) Archipelagic Sea Lane
d) Bridge
e) Island Group
f) Range system
g) Two Way Route

The only difficulties in automatically converting such groups of components into the individual aggregation features are :

1. Determining which aggregation feature to define from the C_ASSO (this could be denoted in the INFORM of the C_ASSO however)
2. Establishing the desired attribution within the aggregation feature (e.g the name of an Island Group or the Colour of a bridge).

Encoding Guidance.

The two considerations above suggest that C_ASSO features are established for the individual aggregation features desired and a definition of the attribution required is done through configuration of a suitable converter (this is suggested as an enhancement to the existing encoder). Certainly much progress has already been made on this front with the existing encoder and some work can be done in existing S-57 data for definition of island groups, range systems and traffic separation schemes.

An example of Island groups definition from existing US data is shown in the following table. This shows three LNDARE and a C_ASSO (linked via FFPT records) being translated into a named IslandGroup with the three Land Area features.

<table>
<tr>
<td>

```
C_ASSO:
{
  OBJNAM = Los Coronados
  FFPT = 2602C6EC940BE211
  FFPT = 2602C5EC940BE211
  FFPT = 2602C8EC940BE211
  FFPT = 2602BEEC940BE211
}
LNDARE (2602C6EC940BE211):
{
  OBJNAM = North Coronado
}

LNDARE (2602C5EC940BE211):
{
  OBJNAM = Middle Coronado
}

LNDARE (2602C8EC940BE211):
{
}

LNDARE (2602BEEC940BE211):
{
  OBJNAM = South Coronado
}
```

</td>
<td>

```
IslandGroup:
{
  featureName:
  {
    displayName=0
    language=eng
    name=Los Coronados
  }
}
islandAggregation consistsOf
{
    LandArea:
    {
      featureName:
      {
        displayName=0
        language=eng
        name=North Coronado
      }
    }
}
islandAggregation consistsOf
{
    LandArea:
    {
      featureName:
      {
        displayName=0
        language=eng
        name=Middle Coronado
      }
    }
}
islandAggregation consistsOf
{
}
```

</td>
</tr>
</table>

Span fixed / Span opening

Span Opening and Span Fixed features should be straightforward to translate from the S-57 BRIDGE source features. Largely the attribution is the same and translates over but the previous section on Aggregation features should also be taken into account in terms of ensuring that the aggregation feature assumes the desired values.

The following example shows a single S-57 BRIDGE feature being converted into a Span Opening feature and an associated Nautical Information feature by association..

S57 BRIDGE1	S101 SpanOpening1
BRIDGE: **{** **CATBRG = 4** **HORCLR = 24.3** ***SORDAT = 200701*** ***SORIND = US,US,graph,Chart 11352*** **VERCCL = 1.2** **VERCOP = 22.2** **SCAMIN = 499999** **}**	**SpanOpening:** **{** **horizontalClearanceFixed:** **{** **horizontalClearanceValue=24.3** **}** **verticalClearanceClosed:** **{** **verticalClearanceValue=1.2** **}** **verticalClearanceOpen:** **{** **verticalClearanceValue=22.2** **}** **}**

The table below shows a slightly more complex example with the Bridge converted into a single Span Opening feature and an associated Nautical Information information type (and also shows the loss of the OBJNAM value – this should be translated into a name on an aggregation feature which uses the SpanOpening as a component.

S57 BRIDGE1	S101 SpanOpening1
BRIDGE: **{** **CATBRG = 1** **INFORM = Railway bridge** **OBJNAM = Huey Point Long Fixed Bridge** **VERCLR =** ***SCAMIN = 499999*** **}**	**SpanFixed:** **{** **verticalClearanceFixed:** **{** **verticalClearanceValue** **}** **}** **additionalInformation informationProvidedFor** **{** **NauticalInformation:** **{** **information:** **{** **language=eng**

	text=Railway bridge } } }

Encoding Guidance.

Overall the only issues with Bridge conversions are to ensure that the converter being used is able to correctly convert and use existing S-57 aggregations to generate the bridge aggregation feature and attribute it based on the individual components. Features that are part of the Bridge, rather than the spans may have to be defined by the individual components (e.g Name, Colour etc) and a review of the existence and attribution of the components will ensure correct attribution.

Wind Turbines

Wind Turbines should be transformed from LNDMRK,CATLMK=19 features in S-57. Sometimes the S-57 feature has a dual function as in the case below where the S-57 LNDMRK has FUNCTN=33 (light support) as well as being a Windmotor (CATLMK=19).

e.g

LNDMRK: { CATLMK = 19 CONVIS = 1 FUNCTN = 33 INFORM = Maintained by Deepwater Wind RI. OBJNAM = Block Island Wind Farm WTG-2 STATUS = 8 }	**Landmark: { visuallyConspicuous=1 function=33 featureName: { displayName=0 language=eng name=Block Island Wind Farm WTG-2 } reportedDate=20151208 status=8 scaleMinimum=89999 inTheWater=1 } additionalInformation informationProvidedFor { NauticalInformation: { information: { language=eng text=Maintained by Deepwater Wind RI. } } }**

If the LNDMRK is encoded with CATLMK=19 then it should be translated to a Wind Turbine. If FUNCTN=33 (light support) then the associated light can be inserted as well. In general the advice to the encoder is

1. Ensure that the function of the feature with CATLMK=19 is unambiguous and only a wind turbine (if another function is also present then it may require its attributes to be inspected to ensure information is not lost).
2. Wind turbines are point features. S-101, like S-57 has different features for wind farms as opposed to individual wind turbines.
3. Additionally it should be encoded so that elements are not lost between the attribute bindings of a Landmark and those of a Wind Turbine, the most notable example being VERLEN which should be translated to Height in the S-101 feature.

Discoloured Water

The current UOC guidance for discoloured water is shown below:

- The text "Discoloured water" on the source indicates the probable existence of shallow water. This should be encoded using a **CTNARE** object with attribute INFORM or TXTDSC containing a cautionary note (see clause 6.6).

Figure 10: UOC Section 6.5 - discoloured water

It should be possible for the converter to find CTNARE objects in the S-57 cell with INFORM containing the text “Discoloured Water” and convert them straight to Discoloured Water features as per the mechanism proposed in Section 2.

In this case the ECDIS behaviour will also need to be defined (as CTNARE currently precipitates alarm/indications so “Discoloured Water” probably should as well? If so, the coincident CTNARE could be removed (if no other CATRES is present)

The advice to encoders in respect of these features is to understand what INFORM text is converted by the converter and ensure all CTNARE features which are destined to be transformed into Discoloured Water are suitably attributed.

Collision Regulations Limit

The definition of Collision regulations Limit features is similar to the situation with Discoloured Water. The UOC currently provides guidance, reproduced below:

Some nations have introduced collision regulations (COLREG’s) that may include demarcation lines differentiating between inland water rules and International Rules as a result of the Convention on the International Regulations for Preventing Collisions at Sea 1972. If it is required to encode COLREG’s, it should be done using a narrow **CTNARE** object of type area (see clause 6.6) covering the demarcation line, with attribute INFORM and/or TXTDSC containing a short explanation about the regulation, (e.g. cautionary note from the paper chart). The attribute TXTDSC may be used instead of INFORM, or for longer explanations or notes.

Figure 11: Current guidance on encoding COLREGs

The obvious strategy for creation of these features is to ensure that they are created from CTNARE features where the INFORM contains "COLREGS" or "Collision Regulations". As per discoloured water the current converter does not create these features.

The advice to encoders would be to ensure correct formatting of INFORM for the citation. Population of the S-101 attribute "Regulation Citation" may require further parsing either of INFORM or TXTDSC in the source S-57 feature.

The other issue with Collision Regulations limit is that the S-101 primitive is a curve but the recommended primitive in the UOC is Area. To account for this the converter could be configured to convert the area feature to a curve (by tracing a spine down the thin polygon) and using preset tolerances in the encoding of the area to ensure sufficient accuracy. Given the lack of a generic "boundary" feature this could be an added use case for the introduction of such a feature to delimit outer limits of areas with national/international boundaries and other areas such as MARPOL zones.

3.3. Population of new Feature Catalogue attributes.

This section explores the possibilities for creation of new S-101 attributes and what, if any, encoding review and update would be necessary.

The main categories of mechanism for population are:

1. INFORM mechanisms whether attribution is already in place by virtue of the UOC encodings or new encodings.
2. Data Quality
3. One time configuration per cell (e.g for quality and/or metadata).
4. Update information.
5. Population via existing attribution in features
6. Assessment and update of individual features

Most new attributes could be populated/defined through INFORM encodings but some are more difficult – in the water is hard because extraneous dataset features may have been encoded because of the ECDIS display issues and these features may require deletion once in the water is populated (a suggested enhancement to automatically populate in the water is suggested).

A full list of new attributes together with comments on possibilities for their creation are listed in the following table:

Attribute	Comment

Category of offshore production area	New CATPRA values for offshore production. Can be parsed from INFORM values (UOC guidance)
Category of preference	Only used in PILBOP translation. Could come from INFORM values. Not currently captured.
Category of Temporal Variation	Data quality attribution only.
Data assessment	Data quality only
Date disused	Only used in dumping ground. Can be identified by STATUS and set from INFORM
Date fixed	Only used in Non-Standard Working Day
Date variable	Only used in Non-Standard Working Day
Day of Week	Only used in Service hours
Day of week is range	Only used in Service hours
Default clearance depth	See 30.1 Portrayal Attribute. Auto-generated
Display Name	Should be defaulted and overridden by INFORM if necessary?
Distance Unit of Measurement	Only used in distance Mark Measured Distance value complex attribute. Could be populated through INFORM
Dredged Date	Only used in Dredged area. Could be populated from INFORM
Flare angle	See 30.1 Portrayal Attribute. Auto-generated (by converter?)
Flare Stack	Only used in offshore platforms. Could be populated from existing coincident features.
Flip bearing	Used in Text placement only.
Full seafloor coverage achieved	Used in quality of bathy survey and quality of survey.
Horizontal Clearance Length	Used in Berths, dry docks, floating dock and dock areas. Could be populated from INFORM.
In dispute	
In the water	Guarantees ECDIS display in BASE. Would need review of features where required (from UOC map) and population from INFORM. Coincident features could then be removed.
is MRCC	Coastguard station only. Could be populated from INFORM
language	Used in names
Least Depth of detected features measured	Only used in quality of survey and quality of bathy survey.
Line spacing maximum	Used in Quality of Survey
Magnetic anomaly value minimum	Minimum values currently in INFORM. Only in magnetic anomaly.
Major light	Could be populated from INFORM. Would need detailed review of all lights to establish status.
Maximum Display Scale	
Maximum permitted draught	Currently populated in INFORM for some features. Used in Berths, dry dock, floating dock, dredged area, obstruction, recommended track, range system, fairway systems, two way routes and traffic separation schemes. Not mandatory.
Measured distance	Only used in navigation lines. Currently populated in INFORM.
Measurement distance maximum	Only used in navigation lines
Measurement distance minimum	Only used in navigation lines
Minimum display scale	Set in Data coverage by converter.
MMSI Code	Used in physical/virtual AIS aids to navigation and contact details. Could be populated from INFORM.
Moire effect	Should be populated from CATLIT=16

Multiplicity known	Used in Building, Conveyor, Cable Overhead, Pipeline Overhead, Pylon/Bridge Support, Landmark, Silo/Tank, Wind Turbine, Pipeline Submarine, Lights. Could be populated from INFORM where necessary. Currenty INFORM=”more than one”
name	From OBJNAM with suitable defaults for language. Part of “Feature Name”
Number of features	See Multiplicity known.
Orientation uncertainty	Only used in Quality of Bathymetric data
Pilot movement	Only used in PILBOP translation. Could come from INFORM. Not currently captured
Reference Location	Only used in distance mark Measured Distance Value (complex attribute). Can be populated from INFORM.
Regulation Citation	Only used in Collision regulations limit. INFORM
Reported Date	Used in many features. Should be a direct translation from SORDAT
Sector bearing	From SECTR1 in S-57 (Check)
Sector Extension	See 30.1 Portrayal Attribute. Auto-generated (by converter?)
significant features detected	Used in Quality of Bathymetric Survey and Quality of Survey.
Size of features detected	Used in Quality of Bathymetric Survey and Quality of Survey.
Speed limit	Only in Restricted Areas. Currently in INFORM.
Speed maximum	Only in Restricted Areas. Currently in INFORM.
Speed minimum	Only in Restricted Areas. Currently in INFORM.
Station name	Part of existing TS_TSP, would need to be configured and parsed out.
Station Number	Part of existing TS_TSP, would need to be configured and parsed out.
Surrounding depth	See 30.1 Portrayal Attribute. Auto-generated (by converter?)
Swept Date	Only in Swept area. Could be parsed from SORDAT?
Telecommunication identifier	Used in contact details
Telecommunication Service	Used in contact details
Text	
Text Justification	
Text type	
Time of Day End	Only used in Service hours
Time of Day Start	Only used in Service hours
Uncertainty fixed	
Uncertainty variable factor	
Underlying layer	Part of surface Characteristics complex type and only used in seabed area. It would be complicated but could be done via INFORM.
Vessel Class	Used in speed limit within restricted areas. Could be done via INFORM.
Virtual AIS Aid to Navigation Type	

3.4. Deprecated Primitives and Enumerations.

Feature-spatial primitive bindings which no longer exist in S-101

The current S-101 feature catalogue details various features for which the spatial primitives are more restricted under S-101 than under S-57. These are listed below (deleted ones in **bold**).

Feature	Allowable primitives (S-57)	S-101 Primitives
BRIDGE	**Point**, Line, Area	Curve, Surface
DAMCON	**Point**, Line, Area	Curve, Surface
GRIDIRN	**Point**, Area	Surface
DEPARE	**Line**, Area	Surface
RECTRC	Line, **Area**	Curve
ROADWY	**Point**, Line, Area	Curve, Surface
PIPSOL	**Point**, Line	Curve
TUNNEL	**Point**, Line, Area	Curve, Surface

These were dropped in S-101 mainly because they have no ECDIS display and no significance for navigation. These features remain in many ENCs however – some, Line LNDARE for instance have been removed via mechanisms in S-58.

From that perspective it is reasonable to allow them to be filtered out as part of the S-101 conversion process but review of selected ones may be warranted depending on the ENC encoder's preferences, encoding guidance and individual feature characteristics.

Attribute enumeration bindings which no longer exist.

Of potentially greater significance is the existence of enumerated attributes within the current global ENC portfolio which are prohibited under the bindings set out by the S-101 feature catalogue. One of the improvements in the S-101 feature catalogue is the ability to bind specific attribute values to a particular feature class, not just the attribute (under S-57 this was defined in the UOC). For example although the attribute “status” has the following values under S-101:

Status: IHO Definition: The condition of an object at a given instant in time.

Attribute Type: Enumeration.

1) **Permanent**

   IHO Definition: Intended to last or function indefinitely. (The Concise Oxford Dictionary, 7th Edition).

2) **Occasional**

   IHO Definition: Acting on special occasions; happening irregularly. (The Concise Oxford Dictionary, 7th Edition).

3) **Recommended**

   IHO Definition: Presented as worthy of confidence, acceptance, use, etc. (The Macquarie Dictionary, 1988).

4) **Not in use**

   IHO Definition: Use has ceased, but the facility still exists intact; disused. (Adapted from Defence Geospatial Information Working Group; Feature Data Dictionary Register, 2010).

5) **Periodic/intermittent**

   IHO Definition: Recurring at intervals. (The Concise Oxford Dictionary, 7th Edition).

6) **Reserved**

   IHO Definition: Set apart for some specific use. (Adapted from The Concise Oxford Dictionary, 7th Edition).

7) **Temporary**

   IHO Definition: Meant to last only for a time. (The Concise Oxford Dictionary).

8) **Private**

   IHO Definition: Administered by an individual or corporation, rather than a State or a public body. (Defence Geospatial Information Working Group; Feature Data Dictionary Register, 2010).

9) **Mandatory**

   IHO Definition: Compulsory; enforced. (The Concise Oxford Dictionary, 7th Edition).

11) **Extinguished**

   IHO Definition: No longer lit. (S-57 Edition 3.1, Appendix A – Chapter 2, Page 2.197, November 2000).

12) **Illuminated**

   IHO Definition: Lit by floodlights, strip lights, etc. (S-57 Edition 3.1, Appendix A – Chapter 2, Page 2.198, November 2000).

13) **Historic**

   IHO Definition: Famous in history; of historical interest. (The Concise Oxford Dictionary, 7th Edition).

14) **Public**

   IHO Definition: Belonging to, available to, used or shared by, the community as a whole and not restricted to private use. (Adapted from The New Shorter Oxford English Dictionary, 1993).

15) **Synchronized**

   IHO Definition: Occur at a time, coincide in point of time, be contemporary or simultaneous. (The New Shorter Oxford English Dictionary, 1993).

16) **Watched**

   IHO Definition: Looked at or observed over a period of time especially so as to be aware of any movement or change. (adapted from The New Shorter Oxford English Dictionary, 1993).

17) **Un-watched**

   IHO Definition: Usually automatic in operation, without any permanently-stationed personnel to superintend it. (Adapted from IHO Dictionary – S-32).

18) **Existence doubtful**

   IHO Definition: A feature that has been reported but has not been definitely determined to exist. (Adapted from S-57 Edition 3.1, Appendix A – Chapter 2, Page 2.198, November 2000).

28) **Buoyed**

   IHO Definition: Marked by buoys. (Australian Hydrographic Service).

Remarks:

- No remarks.

However, a landmark for example can only have the following values within the DCEG:

Status	(STATUS)	2 : occasional 4 : not in use 5 : periodic/intermittent 7 : temporary 8 : private 12 : illuminated 13 : historic 14 : public	EN	0,*

Figure 12: Values of status allowed for Landmark by DCEG

In S-57 the mapping of features to attributes and their individual values was embodied in the UOC which also recommended the allowable values for each enumerated attribute on a per feature basis.

In S-101 this is now more rigidly defined and enforceable through automated means using the XML feature catalogue. When the transformation takes place from S-57 to S-101 attribute values outside those allowed in the feature bindings can be trapped by the conversion process and reported. In this sense the DCEG bindings become mandatory, unlike the UOC's more advisory bindings.

There are two possible strategies during the conversion process then:

a. Respect the attribute values within the originating S-57 feature and map the enumerated value to an equivalent within the S-101 feature catalogue, noting any attribution outside the FC. This is, of course, assuming an equivalent attribute exists (in some cases it may not).
b. Trap the error, report and leave feature unattributed with the "illegal" binding. This has the advantage of leaving a dataset in full accordance with the feature catalogue.

The following table is illustrative only and shows a number of "prohibited" attribute enumerations trapped as part of a large scale conversion process.

Total features	**Feature**	**Enumeration (allowable / prohibited under UOC)**	**Attribute**
12720	UnderwaterAwashRock.	**boulder[18],rock[9],**	natureOfSurface
2026	Landmark.	**permanent[1],**	status
1505	BeaconSpecialPurposeGeneral.	**lattice beacon[4],**	beaconShape
1250	Coastline.	**sandy shore[3],stony shore[4],coral reef[9],**	categoryOfCoastline
698	BeaconLateral.	**lattice beacon[4],**	beaconShape
290	CableSubmarine.	**telephone[4],**	categoryOfCable
290	SiloTank.	**permanent[1],**	status
192	Daymark.	**painted[9],**	natureOfConstruction
130	BuoySpecialPurposeGeneral.	**private mark[13],**	categoryOfSpecialPurposeMark
116	BeaconSpecialPurposeGeneral.	**private mark[13],**	categoryOfSpecialPurposeMark
109	River.	permanent[1],	status
84	UnderwaterAwashRock.	**deeper than the range of depth of the surrounding depth area[3],**	expositionOfSounding
55	*Obstruction.*	*foul ground[7]*[7]*,*	*categoryOfObstruction*
55	BuoySpecialPurposeGeneral.	other system[10],	marksNavigationalSystemOf
53	SlopeTopline.	dune[3],hill[4],	categoryOfSlope
22	Building.	**permanent[1],**	status
21	ProductionStorageArea.	**permanent[1],private[8],**	status
18	*Landmark.*	*windmotor[19],*	*categoryOfLandmark*[8]
14	Sounding.	found by laser[7],	techniqueOfVerticalMeasurement

---

[7] OBSTRN,CATOBS=7(foul ground) should be transformed to a "foulGround" feature so in the context of S-101 this feature/attribute combination is valid (in that it has a valid transformation to an S-101 equivalent).

[8] LNDMRK,CATLMK=19 (windmotor) should become a "windTurbine" feature. However, LNDMRK doesn't support WATLEV (which is allowed under windTurbine, as is vertical clearance) so a mechanism (probably via INFORM) may be needed to populate the equivalent windturbine feature adequately.

12	UnderwaterAwashRock.	found by laser[7],	techniqueOfVerticalMeasurement
9	Retroreflector.	**periodic/intermittent[5],**	status
8	CableArea.	**telephone[4],**	categoryOfCable
6	Daymark.	**private mark[13],**	categoryOfSpecialPurposeMark
5	CableOverhead.	**mean lower low water[12],**	verticalDatum
2	Obstruction.	found by laser[7],	techniqueOfVerticalMeasurement
1	**Vegetation.**	**grass[1][9],**	**categoryOfVegetation**
1	BeaconLateral.	other system[10],	marksNavigationalSystemOf
1	BeaconSpecialPurposeGeneral.	other system[10],	marksNavigationalSystemOf
1	BuoySafeWater.	other system[10],	marksNavigationalSystemOf
1	BeaconLateral.	painted[9],	natureOfConstruction
1	Landmark.	painted[9],	natureOfConstruction
1	CableArea.	not in use[4],	status
1	LogPond.	not in use[4],	status

The current S-101 feature catalogue contains much more prescriptive bindings than those described in the S-57 UOC and is the result of much dialogue and discussion within the relevant IHO working groups. There are few, if any, instances where an attribute binding's exclusion from the feature catalogue can have a navigational impact so the lack of these bindings should have no impact on the ENC's end user. An analysis of tables similar to the one presented here shows no ECDIS portrayal impacts from the loss of these combinations and nothing that impacts the ENC's ability to meet the IMO PS within ECDIS. If anything, the S-101 data is the "cleaner" dataset because of the excluded featture/attribute combinations.

There are two main recommendations to note in terms of this report's deliverables (i.e optimising encoding for translation):

Encoding Optimisation

An encoder should aim primarily to stay within the enumeration values stated within the UOC for an attribute bound to a particular feature instance to ensure a cleaner conversion to S-101 data. In addition, validation against S-58 can highlight where inconsistencies in existing data exist.

Once the feature catalogue is fully baselined an assessment can be done of features which will not have their full attribution mapped. The encoder can then make an assessment as to whether this is acceptable in terms of information loss or whether alternative encodings should be considered[10]. From the FC the list of "legal" mappings can be determined and, from a consideration of the encoder's own encoding guidelines and the current ENC portfolio, a query can determine the extent to which either information will be lost or features will need re-attributing. Some combinations may be more significant than others but it should be borne in mind that S-101 is a "product" and S-57 to S-101 conversion is a product-to-product conversion so, although certain attributes are not present in the eventual dataset the encoder

[9]Although allowed by the UOC, VEGATN,CATVEG=1(grass) does not display in ECDIS so is rare in live data.
[10] This pre-supposes option (b) in the previous section – i.e that the decision is taken to only allow "legal" encodings within the destination S-101 dataset

still can maintain them in their individual production system. The impact analysis of the enforced nature of the feature catalogue should focus on any genuine end user impact and all indications are that there is no navigational risk

The preceding table contains an non-exhaustive list based on limited ENC providers and the feature catalogue at the time of writing but it is likely that both ENCs and the FC will progress in the coming months.

3.5. Geometry Conversion.

Geometry Equivalence

It is relatively easy to define geometric equivalency of an S-57 feature or set of features with its S-101 translation.

Every feature includes links to its geometry components. As geometry in S-100/S-57 is indirect it is possible to "explode" the links and realise them to an ordered set of points. Two features can be seen as equivalent (from a pure geometry perspective with no attribution) if they resolve to the same points. Specifically:

- Two points (an S-57 one and an S-101 one) are equivalent if they have the same coordinates.
- An S-101 curve is equivalent to an S-57 set of edges if they resolve to the same set of geometric coordinates in the same order.
- A surface is equivalent to a polygon again, if the realised points are the same and the digitizing direction (which defines the topology) are the same.

So, in order to be systematic about how S-101 transformation takes place it is enough to consider how the two geometry models differ and how to form a one-one link between the two.

S-57	S-101
1. Points are defined as Isolated or connected and by their location. Every point has a unique identifier within the cell 2. Individually named edges are defined by referenced start and end points together with a set of intermediate vertices – interpolation is fixed as "loxodrome" 3. Cell features requiring geometry link to an ordered list of geometry records:	1. Points are defined 2. Named curves are formed from start and end points together with intermediate vertices. – interpolation is fixed as "loxodrome" 3. Curves can themselves be aggregated into composite curves (which can reference other composite curves) 4. Composite curves can be aggregated into surfaces complete with holes where required

a. Point features link to the individual point record required b. Line features link to an ordered list of edges (each of which can be oriented either in the direction in which it is digitised or the reverse) c. Area features link to an ordered list of edges which define the exterior and (possibly multiple) interior holes. As with line features each edge can be individually oriented. 4. It is up to the encoding application (the ENC authoring system) to ensure topological consistency (mainly, the correct labelling of holes in polygons, the existence of all linked geometry components and the consistency of update instructions)	5. Cell features requiring geometry link to an ordered list of geometry records: a. Point features link to the individual point record required b. Curve features link to sets of either curve or composite curve records c. Surface features link to composite curves or surfaces 6. It is up to the encoding application (the ENC authoring system) to ensure topological consistency (mainly, the correct labelling of holes in polygons, the existence of all linked geometry components and the consistency of update instructions)

Figure 13: Comparison of geometry models, S-57 and S-101

The following diagram shows a simple example.

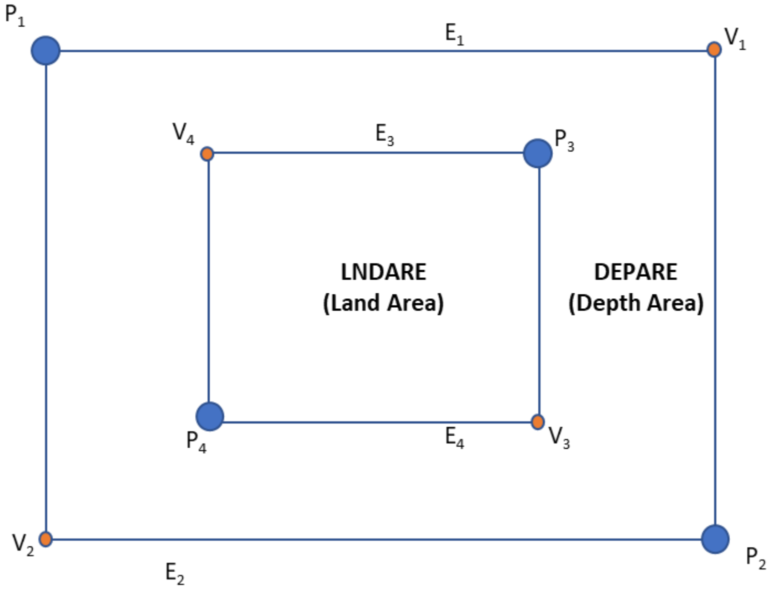


Figure 14: Example: a simple area feature and its geometric resolution.

In this example a single Land Area feature (LNDARE) is surrounded by a Depth Area (DEPARE). The geometry of the Land Area is made up as follows:

Edge 1 = Point 1 x Vertex 1 x Point 2
Edge 2 = Point 2 x Vertex 2 x Point 1
Edge 3 = Point 3 x Vertex 4 x Point 4

Edge 4 = Point 4 x Vertex 3 x Point 3
In the S-57 case the Land Area geometry is made up of a sequence of the two edges E3 and E4 and the Depth Area geometry is made from Edge 1 and Edge 2 with Edge 3 and Edge 4 as an interior hole. In the S-101 case two additional composite curves can be made where:

Composite Curve 1 = E1 x E2
Composite Curve 2 = E3 x E4

And a single surface:

Surface 1 = Composite Curve 1 with a hole Composite curve 2

The Land Area has geometry Composite Curve 1 and the Depth area Surface 1.

So, although the geometry is identical the S-101 model has introduced two higher levels of abstraction to simplify the interface between the features and their geometry components. In practice this makes for great efficiencies when complex geometries are shared by features (e.g Coastline sharing perimeters of depth areas and land areas, quality features covering dataset boundaries etc) without changing the underlying coordinates of the geometry.

Conversion of Geometry between S-57 and S-101

Even though S-101 has an additional layer of geometric abstraction the points the features resolve to are the same. So the display of the features in terms of their position on screen will be the same. There should be no geometric editing of S-57 features needed to optimise the conversion to S-101.

It is easier to quantify the issues with the geometry of S-57 to S-101 conversion. Essentially, there are only two changes to S-101 geometry:

a) The introduction of another layer of abstraction, in the creation of composite curve geometries which are associated from individual features. In S-57, each feature realises its own geometry from the array of defined points and curves, So (still in S-57) if two area features share the same geometry with, say, 15 edge components, then both features would have 15 feature->spatial pointer fields. In S-101 the 15 edge components would map to a single composite curve which would, in turn, be associated with the two feature instances. This extra layer of abstraction helps to cut down needless parallel construction of coincident geometry between features in the same dataset within S-101
b) The changes to group 1 “skin of the earth” (SOE) features. As these features tile the plane geometric transformations are required in order to maintain the integrity of the SOE. These changes will transform the curves, composite curves and spatial pointers within the S-101 dataset.

Encoding optimisation.

In respect of (a) above there is little encoding optimisation which can be (or needs to be) performed. It is possible to form an injective mapping from the S-57 geometry to the equivalent S-101 geometry (i.e everything in S-57 maps to a unique S-101 equivalent (and, by implication, backward conversion back to S-57 geometry is guaranteed). Additionally, the extra composite curve features can be completely defined and optimised by the existing S-57 FSPT links so no extra optimisation is possible (or required)

Skin of the Earth Changes

With (b) the situation is a little more complicated. The SOE feature change requires a spatial operation every time certain features are converted to their S-101 equivalents. The diagram below illustrates the new SOE mappings embodied in the current S-101 DCEG/PS:

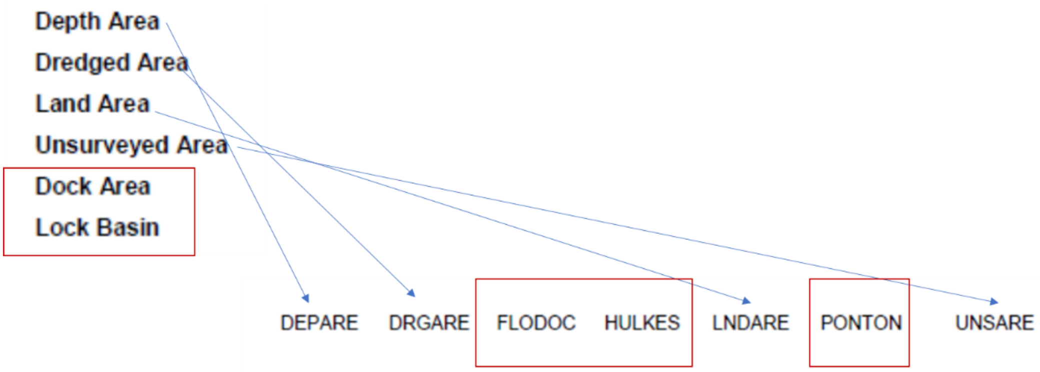


Figure 15: Skin of the Earth Changes between S-57 and S-101]

How to effect changes in geometry with SOE features.

There are two operations that need to be done in order to effect the change to SOE features:

1. The new SOE features must have their geometry inserted into the existing SOE, effectively by "punching" in a hole into the existing SOE.
2. Features which are no longer SOE must have their geometry covered by existing SOE features.

Punching Holes:

To insert new SOE features into the existing (converted S-57) skin of the earth, new features (e.g Dock areas and lock basins) can be "cut out" and inserted directly into the old s-57 group one features. This requires the construction of new coincident edges (and may result in the bisection of existing group 1 features and the need to replicate their attribution. This, however, should be a fully deterministic process and no encoding optimisation is possible or required. Full testing of this feature has not been carried out during the writing of this report.

Filling in Holes:

There are three features which are no longer SOE in S-101, floating dock (FLODOC), Hulks (HULKES) and Pontoons (PONTON). When conversion to S-101 takes place these features (or rather, their equivalents) cease to be part of the skin of the earth and therefore require the "hole" in the dataset's SOE to be filled in. A simplistic approach would be to simply fill in the hole with an unsurveyed area and overlay the equivalent S-101 feature on top (with coincident geometry). This requires little geometry manipulation and is unambiguous. Additionally, should a reverse conversion to S-57 be required this could be accomplished by identifying the feature and reversing the process (effectively absorbing the feature back into the Unsurveyed area. The diagram below illustrates this process (as viewed in 2-D "from the side" where the SOE goes across from left to right).

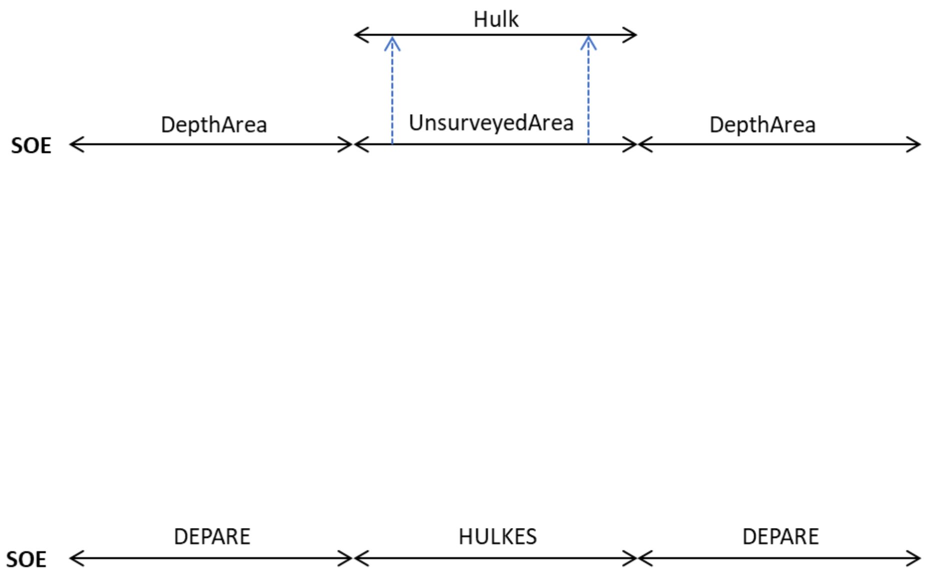


Figure 16: Adjusting the skin of the earth

As can be seen from the preceding diagram the HULKES$_{S57}$ feature is converted into two features with coincident geometry – an UnsurveyedArea$_{S101}$ and a Hulk$_{S101}$. The depth areas either side (making up the seamless skin of the earth) remain and border the new SOE UnsurveyedArea feature. E.g:

^{S57}**HULKES**$_1$	^{S101}**Hulk**$_1$ **+** ^{S101}**UnsurveyedArea**$_2$
`HULKES:` `{` `  CATHLK = 1` `}` `geometry` `{` `   G` `}`	`Hulk:` `{` `  CategoryOfHulk = 1` `  geometry:` `  {` `    g` `  }` `}` `UnsurveyedArea:` `{` `  geometry:` `  {` `    g` `  }`

	}

Other possibilities:
It would certainly be possible to substitute other features instead of UnsurveyedArea in place of HULKES, FLODOC and PONTON features during the conversion process but the choice of feature(s) and their attribution may not be straightforward. Certainly in the following case:

Figure 17: SOE example2, a FLODOC feature next to other group1 ENC features

In the example it would be tempting to continue the depth area borders across the hole left by the FLODOC's removal but this would need some cartographic consideration and would be difficult to carry out automatically in the general case. In the illustrated case the hole left by the FLODOC could be defined as a Depth Area and take on the shoalest DRVAL1 from the surrounding depth areas instead of its replacement with an unsurveyed area. This is an area where configurability of the converter is crucial.

Reversal of conversion S-101 -> S-57…

An interesting point raised by the geometry considerations is whether the reverse conversion could be achieved (i.e from S-101's level 3 geometry to S-57s chain node topology). The basic geometry would certainly be possible and by replacing the SOE features as documented the process could be reversed. This is not considered in detail in this document but would be a good area for further study in connection with future distribution models.

Conclusions and Encoding advice.

In terms of encoding of S-57 data it is probably best to assess all Skin of the Earth Changes (i.e examine features where their SOE status changes), ensure the surrounding features do not give cause for ambiguity and ensure the conversion process will behave in a known way when doing the conversion.

4. Further Work

The following were summarised earlier as suggestions for further work following this report which may be of interest to the community specifying S-101

a. Conversion Specification
b. Distribution and conversion methodologies
c. Converter enhancement requirements
d. Test data in a particular area
e. Updates

4.1. Conversion Specification

In producing this report it was difficult, at times, to understand how a particular S-57 object should be translated into an S-101 feature. A great proportion of the S-57 feature catalogue can be translated directly with the S-57 acronym mapping to the S-101 alias embedded in the S-101 feature catalogue. However, there are more complex transformations, e.g

1. Where an S-57 object should be translated to a different feature because of its attributge values (e.g creation of Foul Ground features).
2. Where an S-57 object creates another feature during the process of its conversion (e.g information types created from INFORM content or creation of features during Skin of the Earth integration)
3. Where features combine to form a single S-101 feature or where aggregations need to be converted (e.g sectored LIGHTS or C_ASSO to Bridge Aggregations)

Because of these complexities a potential further piece of work would be to specify a way of describing translation of S-57 cells to S-101 datasets which is independent of the implementation. This would allow the intention of the DCEG writers and the S-101 community to be expressed for the creation of new S-101 datasets.

Such a specification should not be compex and would require many tables of values to be assembled which map the S-57 feature catalogue to corresponding structures in the S-101 feature catalogue (and would probably necessitate the creation of a formal feature catalogue for S-57). The language of conversion is not as simple as straight substitution though (as described in the examples within this document) so a rich language for describing the conversion process would need to be defined. The specification should be machine-readable however so that conversion from S-57 to S-101 can be validated – this would allow a more concrete definition of "success" (as defined in the introductory section to this report) to be defined.

4.2. Conversion and Distribution Models

In order to formulate a structured way forward is worth considering what the transition and ultimate goal of S-57->S-101 conversion is. Although a tool has been produced which is capable of producing S-101 which is largely compliant with the criteria for success, various elements of this report show that consideration should be given to the options for conversion within various distribution paradigms.

Although the ultimate goal is, of course, S-101 becoming the adopted standard for primary SOLAS navigation and S-101 meeting the objectives of the IMO performance standard, there is a long way to go before that goal is reached and, likely, a long interim period where end users are likely to have a combination of S-57 and S-101 capabilities and where producers are transitioning between the two models. So, it is highly likely that a substantial transition period should be part of the S-101 rollout plans. In simplistic terms the situation is as illustrated in the diagram below:

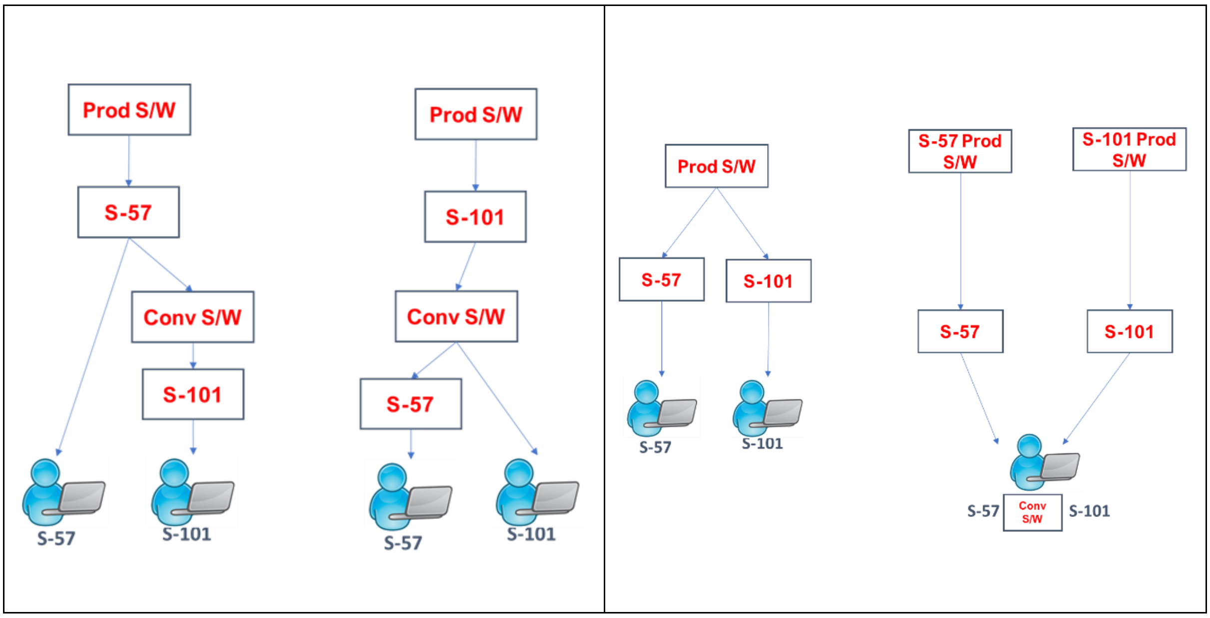


Figure 18: Options for S-57 / S-101 conversion and distribution

This study looks at optimising the encoding of S-57 ENC data in order to optimise a subsequent conversion to S-101. It is helpful to consider multiple use cases for S-57 to S-101 conversion as they may fit individual member state preferences in the future. In the diagram above various options for the ENC producer (the member state) and the path of data to an end user are presented. The options described are, roughly:

1. The member state produces S-57 ENC data from their production systems, encodes it as S-57 and then converts it into S-101 for distribution to end users. In this context the

data is “optimised” in its S-57 encoding so that when it is converted to S-101 data it works as well as possible. The full S-101 functionality may not be available to be used in the S-101 data but there should be an expectation of the S-101 data providing an enhanced user experience for the end user.
2. The member state produces S-101 ENC data from their production systems, encodes it as S-101 and then converts it into S-57 for distribution to end users who are still on legacy S-57 systems. In this context the member states’ main concern is that the legacy S-57 data works in an “equivalent” way to the full S-101 data within the end users’ system and that, although not all S-101 features are present, a guaranteed minimum level of functionality is available.
3. The member state produces both S-57 and S-101 data from a single superset of data. In this context although no “conversion” is done there is still a need (in the eyes of the member state) for “equivalence” between the two encodings of the data, i.e assuring them that the end user experience has a minimum level of (safe) working and that the data content is “equivalent” in its content.
4. Member states produce either S-57 or S-101 data and the end user system converts data as necessary on receipt.

Wherever conversion ultimately takes place it is clear that from a member state point of view is that there is likely to be a period of “parallel running” for some (if not all) end users and the question, then, of functional equivalence of ENC data in both forms within client systems should be addressed. Functional equivalence in this context means, effectively, that an S-57 ENC and its equivalent S-101 ENC should meet the minimum performance standard set out by IMO in terms of the presentation and functionality in respect of the data content (similar to the criteria for success outlined in this document). The S-101 data can (and we hope should) present the user with a better user experience (fewer extraneous alarms, more intuitive and structured data interrogation and more dynamic updates to feature/portrayal catalogues) but certainly there should be a minimum performance which both datasets conform to which can then be used to assess whether a conversion process has been successful.

Further work in this area would also need to take into account the question of updates and how they can be produced, converted and distributed, whether the reverse conversion of S-101 to S-57 is possible and how that should be approached.

The technical aspects of this process should be laid out as they form the basis for all options in rolling out S-101.

4.3. Converter Enhancements

Further work on converter specifications is proposed as a further piece of work. One of the conclusions of this report is the observation that the more sophisticated the converter technology, the more the new features (and attributes) of S-101 can be defined from within existing S-57 and the more “successful” the conversion of the S-57 data to S-101.

This would entail a more detailed specification of the converter enhancements required to carry out many of the encoding optimisations proposed in this report as well as a more systematic investigation of current technology in this area based on structured test datasets. This would also provide a basis for testing an independent specification of conversion as proposed.

The main proposals for enhancing the existing conversion utility would be:

1. INFORM attribute parsing for population of S-101 features
2. Configuration for automatic classification of major lights both by attribute value (e.g nominal range) but also be configuration of the conversion utility.
3. Aggregation conversion to new features – ensuring that all aggregation features and their attributes can be specified within the component features (this would need rigorous testing by structured test data)
4. Exhaustive testing of new features, attributes and associations along with edge cases and metadata specifications.
5. Metadata configuration for M_QUAL, M_COVR via configuration files.
6. Auto-generation of inTheWater attribution.
7. Configuration to support richer transformation of skin of the earth features. (e.g substitution and intelligent creation of new SOE features and replacement of non glacier ICEARE underlying SOE feature unsurveyed area with land area)
8. Creation of new primitives (creation of COLREGS from areas, creation of Piles areas/curves from individual points)

4.4. Test Area of conversion

A worked example within a particular area based on a single member state (e.g US data) and using a genuine encoding guide to focus reviews of existing data and its conversion to S-101. This would be a detailed assessment of some of the areas defined in this document based on real data and would provide a good benchmark for the effort required to produce b9oth a minimum standard of "success" and a fully optimised set of S-101 datasets

This would also necessitate the creation of complete datasets complete with all catalogue and discovery metadata and would help to test an independent specification for conversion (which would form part of a validation suite for the created dataset).

This piece of work can also showcase features, "before and after" conversion and show how the mechanisms of conversion work. The production of test datasets to support conversion testing would also be an integral part of the effort, as would enhanced converter functionality designed to optimize S-101 creation.

4.5. Updates.

Although the question of updates could well be answered as part of a larger piece of work on distribution methodologies the technical aspects of producing updates for S-101 data and whether it is possible to produce them by conversion could also be approached as an independent topic. Certainly it is vital that updates form part of the S-101 methodology for the future and their production, conversion and validation is still at an early stage. It would be useful to answer the question of how they can be produced, converted and applied and to then feed this into discussions of future distribution model