CHRIS15-5.3A

15th CHRIS MEETING
IHB, Monaco, 10-13 June 2003

Summary of study report

'INTEROPERABILITY THROUGH HYDROGRAPHIC STANDARDS HARMONISATION'

by UKHO from a Report by IDON Technologies Inc.

Summary and Proposal *(by IHB)*

Executive summary:	This paper analyses and compares the standards S-57 and DIGEST , as well as their respective based products ENC and DNC, and provides recommendations to harmonize both standards on the occasion of ISO profiling for the next generation, with a view to achieving interoperability between ENCs and DNCs.
Actions to be taken:	The CHRIS committee is invited to consider adding a new CHRIS work item to study this report and its recommendations, and assess if changes to the CHRIS Work Plan are needed, e.g. instructions to TSMAD for the ed 4.0 exercise.
Related documents:	• S-57 Edition 3.1 published in 2000 • DIGEST Edition 2.1 published in 2000

Fundamentals

At the level of the base standards:	ISO profiling for the next generation of each standard should provide compatibility provided the work is done in co-operation and wherever possible the same profiles adopted.
At the product specification level:	There are currently significant differences such that inter-conversion is only possible with manual intervention and will still always result in information loss or degradation. The use of XML based auxiliary tables should allow clean conversion within a production environment. Many of the differences stem from different overall design aims for ENC and DNC. Equally, AML and TOD have very different aims; TOD supports submarine operations and is tailored strongly toward that mission while AML has a much broader scope. The main aim must be to achieve interoperability from the *user* perspective; dual-fuel ECDIS systems have therefore done much to alleviate the problems perceived even a few years ago. They are however, more a sticking-plaster than a cure-all or silver bullet. Migrating the 3000 ENC cells and 2000 DNC libraries currently produced will not be trivial. Any changes to fielded data are more likely to be later rather than sooner.
At a data capture level:	Many of the *perceived* differences between ENC and DNC stem from differing capture criteria and this has often camouflaged the more serious, structural differences.

Scope

The report has been compiled by IDON Technologies Inc of Ottawa, Canada under a contract let by the UKHO in October 2002 and which is jointly funded by the UKHO and NIMA and which stems from action items accepted by these organisations at the DGIWG Steering Committee in Paris in 2001. **This summary has been prepared by the UKHO based on the 12 March 2003 draft of the full report and will be updated in line with the final text when that is available.** The terms of reference for the study outline its purpose as being:

> *"The requirement is to establish an independent assessment of the interoperability requirements for digital hydrographic products, particularly ENC/DNC and AML/TOD; to assess the current level of interoperability; to identify the interoperability problems and their implications for users; to propose a plan for the resolution of those problems with sufficient estimates for changes required in production systems, legacy product data, and end-user systems so as to allow all stakeholders to identify the way forward and then produce outline cost estimates."*

High Level Aspects

The report looks at the issues at three fundamental levels:

- Base exchange standards – S-57 Edition 3.0 was published in 1996 with a minor revision (3.1) produced in 2000; DIGEST 2.1 was published in 2000 though it must be noted that virtually all DIGEST vector data is in practice compliant to the US Mil-Spec 2407 VPF published in 1996.

- Product Specifications – the ENC product specification, is currently published as an appendix to the overall S-57 standard; DNC, though the latest version of the specification was published in 1997 is in practice essentially based on the version of VPF current when it was first written in 1992.
- Product implementations including data capture specifications.

S-57 – DIGEST

There has been a long history of the alignment of DIGEST and S-57 even though these two specifications evolved from two complementary efforts. DIGEST emerged from military land mapping activities, which very early recognised the need for standard digital products to support military operations. Similarly the international hydrographic community, which had standardised the format and presentation of the nautical chart, also saw the need and importance of international standards for hydrographic information.

The report reviews the two standards in terms of the current priority to improve interoperability of products based upon these standards, looks at how the standards are evolving, and considers their relationship to the ISO TC211 work.

Development of each began in early 1980's, initially to address the producer to producer exchange requirements of their respective communities. Since the mid-1990s there have been various efforts to bring about greater alignment between them. DGIWG and IHO are both now looking to recast the next versions of their standards as profiles of the set of, largely abstract, standards being developed by ISO TC211.

The report outlines the existing structures of both DIGEST and S-57, noting the way in which existing elements will require mapping to the various TC211 standards within the next generation of standards. (Part II, Section 2 and 3).

The report also summaries the scope and interrelationship of the (currently over thirty) TC211 standards (Part II, section 4 and 5)

The Open GIS Consortium (OGC) is an international industrial consortium (both NIMA and UKHO are members) which is focusing upon providing specifications that support the delivery of practical systems or components of systems that can work together and facilitate a wide variety of functions for finding, accessing, integrating and sharing of data; that is, they provide for data interoperability. (Part II, Section 6)

The DGIWG and S-57 standards should be built upon the formal ISO standards, but it is desirable to also build them in alignment with the OGC implementation specifications, because this will mean that low cost Standards-based Commercial Off The Shelf (SCOTS) solutions exist.

ENC – DNC

The Electronic Nautical Chart (ENC) product from the International Hydrographic Organization (IHO) and the Digital Nautical Chart (DNC) product co-produced between the United States National Imagery and Mapping Agency (NIMA) and some North Atlantic Treaty Organisation (NATO) allies are two very different products. Although they both derive from the same original paper chart, and ostensibly carry the same data, they have organised that data quite differently and there are significant differences at the detailed level between the two products. This is not to say that the two products cannot inter-work, or be produced from the same production stream, or be updated together. But they are not identical products in different clothing.

Both the products are based on generic standards, the IHO S-57 exchange standard, and the DGIWG DIGEST exchange standard Vector Relational/Vector Product Format. However, both products are defined by product specifications that detail the selected data structures and data fields taken from the generic standards. They also both select from a data dictionary of feature and attribute definitions and bind these features and attributes within the data structures established in the product specification. Both also have associated portrayal specifications, IHO S-52 and NIMA GeoSym.

The parallels between the two products are great, but they differ because the products are intended for somewhat different purposes. They also differ at a more mundane level because the different groups that developed them made a number of different arbitrary low-level decisions.

<u>Specification Differences</u>

According to IDON's understanding the fundamental differences are that the original DNC's purpose was to capture paper chart information and to provide a digital equivalent, while the ENC was designed to drive the IHO defined Electronic Chart Display and Information System (ECDIS).

One result of this is that, at the Product Specification level, ENC is richer in terms of:

- Amount of quality and other metadata required
- Structuring of information to both associate all components of a logical structure such as a traffic separation scheme and also to optimise processing of the information by the ECDIS

However, DNC mandates that all information from the paper chart is captured; ENC allows information deemed to not be of navigational relevance to be omitted. Some ENCs will therefore match DNC content, others will appear thinner. If complete interchangeability is to be achieved then the capture specification for these latter cells would need to be enhanced accordingly.

There is a need to determine how the ENC and DNC can be merged at the data content level so that the default is the more rigorous data requirement. It may be that ENC needs to support more rigorous data capture requirements than just movement of deep draft ships (environmental protection, port security, etc.) while DNC needs to adopt many of the data constructs of the ENC.

ENC should look at raising its topology level at least during production and validation to that of the DNC if it wants to be robust enough to support these newer missions. The production database on the other hand must be validated at full topology (Level 3).

The major specification differences can be summarised under a number of headings and the report describes each in detail with suitable examples. (see Part III Section 3)

- Accuracy, quality and reliability information – ENC is able to carry far more information regarding, for instance, survey quality and reliability through its use of meta objects
- Topology and geometry – DNC only maintains topology within each of its twelve thematic layers, differences in what should be the same geometry for features in different layers can therefore be included; equally, ENC does not enforce the requirement for Group 2 boundaries to be properly shared where the true co-ordinates match
- Bathymetry – especially foreshore; it can be difficult to correctly model depth areas in DNC mainly due to its separating the necessary information across two coverages – hydrography and earth cover.
- Handling of soundings – where they share the same metadata, ENC can carry many soundings as a single '3D' feature; all soundings in DNC are an isolated node with a depth value attribute.
- Cartographic Data – both text features and unstructured, display-oriented attribute values are used in DNC while ENC makes much greater use of name attributes and enumerated attribute values that can either be processed or parsed for display.
- Feature object relationships – where a single logical entity such as a traffic separation scheme, is composed of many simple features, ENC provides a mechanism to include a feature that identifies the set as forming the complex entity.
- Attribute mapping – In a number of cases either attributes in each standard do not use encoded values that map to each other or one standard may use an attribute values for what the other carries by means of a separate object.

- Facility to include repeating attribute values; ENC allows certain attributes to carry repeating values so that, for example all the colours of a light structure can be encoded. This facility is included within DIGEST VRF but is not included in DNC.
- Handling of external files – ENC includes an attribute that allows links between features and external text and/or image files to be established and can include these files within the overall exchange set
- Geometric inconsistencies in DNC – inability to have coincident entity node (point) and connected node. This is permissible in standard VPF but is excluded by the DNC implementation. This, and certain other limitations have also been included within AML and TOD specifications and these would equally benefit from use of the modern form of the standard
- The limitation of DNC to use of the English language only while ENC supports multiple languages

Updating which used to be a major differential between the two products is now less of an issue. NIMA have successfully implemented their VPF Database Update (VDU) system which utilises a binary patch technique similar to that used by major software distributors. In the medium term however, they will move towards system of library replacement with the latest products being downloaded on an as needed basis.

<u>Integrated Content Model</u>

The purpose of this report is not to change the existing ENC or DNC products, but rather to allow both ENC and DNC to be:

- produced from common source;
- updated from common update information;
- used as source for each other (translated into one another);
- used as a common base for additional information such as Additional Military Layers (AMLs).

To do this, it is proposed that one could develop a common integrated content model for both ENC and DNC for the representation of both sets of data within a production system or within a database. The content model would carry the superset of all of the features and attributes allowed for each product. These features and attributes would be described generically independent from either the FACC or IHO Object Catalogue. Also the organisational differences would be handled by an extensive use of collection objects (as defined in S-57). The collection object structure is more flexible than the group approach as it allows both the ENC and DNC data organisations to coexist with the same database, and be recoverable. (Part III, Section 4).

ENC/DNC Implementation

A number of definitions and descriptions of interoperability are provided, perhaps the most relevant is that from the US Joint Staff publication,

> *"the ability of systems, units or forces to provide and receive services from other systems, units or forces, and to use the services so interchanged to enable them to operate effectively together. The conditions achieved among communications-electronics equipment when information or services can be exchanged directly and satisfactorily between them and/or their users."*

A number of types of interaction are identified and analysed (Part IV, Section 1):

- Mutual updating – due to the difference in technology this would, in practice, only be applicable for application of ENC updates to a DNC
- Product to product conversion – currently direct conversion will lose information and requires operator interaction. The report proposes the specification and use of a series of

XML based auxiliary tables to carry all the information nor present in the 'other' product and thereby allow producer based conversions It describes 28 differences between the current products - essentially considering the impact of the specification differences on product to product (including 'round trip') conversion

- Common production – the current moves in both UKHO and NIMA towards production systems utilising feature-based integrated databases should allow for this in the future where necessary. Both ENC and DNC should be produced from the production databases but due to their differing missions and working environments won't necessarily be sub/super-sets of each other. Feature IDs could be implemented in the production databases even though they may not necessarily be immediately incorporated into the products.
- Dual fuel systems are seen as a way to 'buy time' for the next generation of standards based on common profiles of TC211 and authoritative referencing between registers.

It is reported by NIMA that changes to fielded data (DNC and TOD) are more likely to be later rather than sooner.

Achieving Interoperability

Interoperability of the ENC and DNC products is feasible (Part IV, Sections 3 and 4). There currently exists an underlying compatibility inherited from the paper source from which the two products are derived.

The "dual fuel" ECDIS system allows for a transition strategy that buys time for closer integration of the ENC and DNC products. Since a dual or multi fuel capability is a requirement for both the RN and the USN this appears to be a very viable approach.

Direct conversion from ENC to DNC and vice versa will always lose information. An alternative strategy is conversion into a common integrated database that supports a superset of the ENC and DNC requirements. Auxiliary tables should be developed to contain any additional information and relationships so that conversions are repeatable[1]. Auxiliary tables are used to provide the sum of all the information so that there is minimum information loss.

The S-57 object catalogue data dictionary and the FACC data dictionary need to be represented in terms of a registry to support authoritative referencing. The ISO TC211 standard on Registries ISO 19135 and the profile for FACC data dictionary ISO 19126 can be used as a structure for such registries. Registries are simply formalised managed lists that contain unique IDs and metadata about the entries. Authoritative referencing allows registries to cross reference elements from one register to another in a formalised manner. Cross-referencing between registries will identify the matching features and attributes.

DGIWG should reference the IHO object catalogue for the hydrographic features and IHO should reference a number of the FACC features. This will eliminate many inconsistencies; in particular all of the 1 to 1 mappings can be made consistent. This is probably the major alignment that can be made.

Collection objects should be used to aggregate the part of the many to one and one to many relationships, and to allow for the addition of attributes to feature objects. These collection objects should become part of the integrated database and may be expressed in auxiliary tables to the distributed products. Combined with the alignment of the 1 to 1 mappings this will address the majority of problems.

Agreement needs to be achieved on the modeling of the relationship of the parts of certain important multi-part objects, in particular traffic separation schemes, lights and buoys. For these three examples the S-57 structures should be taken as the base, but the FACC components should be fitted into the S-57 framework.

[1] For example in a conversion from a DNC to a ENC, it is proposed that the original DNC data set would result if the process were reversed

Interoperability is feasible, but only if the two driving groups work together. They must harmonize what they have now, and then work together to converge. Otherwise there will be a gradual divergence, and interoperability will not be achieved. The difficulty is that neither existing product can change because they are deployed. All that can be done is to add auxiliary information and align data dictionaries. This will ease common production and multi format display systems such as "dual fuel" ECDIS systems, and to a certain extent ease conversions. More importantly it will create a tendency toward conversion that will emerge in the next editions of both products.

AML/TOD

AML and TOD are both series of products designed to provide information beyond the standard navigation chart (Part V). Though ostensibly similar, there are a number of significant differences.

TOD is a US-national extension to DNC and is strongly tailored to the support of submarine operations; each product in the series essentially equating to one existing paper chart product. Many TOD products would accordingly only be available to the US military. All TOD products use the VPF data format and are only to be used with the equivalent DNC library. The TOD products whose specifications were included in the review were:

- TOD0 – equivalent of OPAREA and NAVEX charts
- TOD1 – equivalent of Bottom Contour charts
- TOD2 – equivalent of Bathymetric Navigation Planning Charts
- TOD4 – equivalent of Hull Integrity Test Site Charts

The AML specifications considered were:

- RAL – Routes Areas and Limits
- CLB – Contour Line Bathymetry
- ESB – Environment, Seabed and Beach
- LBO – Large Bottom Objects
- SBO – Small Bottom Objects
- MFF – Maritime Foundation and Facilities

AML has been developed under the auspices of the NATO Geographic Conference’s Ad Hoc Hydrographic Working Group and is intended to replace the full range of NATO non-navigational hydrographic products. AML therefore covers a far wider range of information types than does TOD, having some ten times as many feature classes within the set of vector products.

Given this different focus and purpose there is little likelihood of merging the two series in the foreseeable future.

Conclusions and Recommendations

Interoperability of the ENC and DNC products is feasible; however, neither carries all the information required to generate the other product.

Over the years, progress towards interoperability has occurred before:

- Since this problem was first discussed perhaps ten years ago, we have seen the same fundamental data models adopted by both standards and far closer agreement of feature and attribute definitions.
- Adoption of ISO TC211 base standards will significantly improve the next generation of both standards and the opportunity must be taken to develop these in co-ordination and wherever possible, adopt identical profiles.
- Dual fuel ECDIS platforms are now common

However, it is equally clear from this new work that no straightforward solution exists; considerable effort will be required to ensure a commonality of approach for the next generation of standards and as well as development of standards, there needs to be a commitment that the production implementations will also be modified to adopt such changes.

Harmonisation MUST be at the production database level and not at the product level, as must exchange and maintenance.

The report concludes (Part VI) with a set of 44 recommendations (split into Long-term and Short-term) which are summarised in the attached table and cover seven major areas:

- Data Models
- Data Dictionary
- Presentation
- Methods
- Exchange of data
- Conversion
- Testing

Way Ahead

- Ensure common data models – much of this is already underway with TSMAD representatives joining the relevant DGIWG project teams.
- Adopt a common structure for data dictionary registers with authoritative referencing between them. Again, an area already recognised by both DGIWG and IHO with essentially identical proposals on the table for their respective next generation of standards.
- Move towards common presentation rules – though both S-52 and GeoSym include essentially the same basic symbols, certain of the rules are handled differently. Both groups are looking to define ISO 19117 compliant rule sets for their next generation and are aware of the need to do so together.
- Investigate how the ENC and DNC can be merged at the data content level so that the default is the more rigorous data requirement; this may include the requirement for ENC to support more rigorous data capture requirements while DNC needs to adopt many of the data constructs of the ENC.
- In the short term, adopt XML based auxiliary tables to carry the information necessary for full (and reversible) interconversion.
- Even though VPF and S-57 are retained as exchange formats, both organisations should also look to the adoption of an XML based encoding for the actual data sets and both will be looking to the work of ISO 19118 encoding standard and ISO 19136 covering GML.
- Enhance testing regimes to ensure both that what ought to be tested is tested thoroughly, but also that ‘false errors’ do not impede product conversion and joint production.
- The emphasis should be on the goal of invisibility of the differences to the operator -on the ship’s bridge or elsewhere - not how it has been partially accomplished so far.

Annex A – Summary of Recommendations

RECOMMENDATION	DESCRIPTION
DATA MODELS	
L1 Data Model	Develop common (DGIWG/IHO) Data Model for the next generation of the two standards Based on ISO abstract standards. Much of this work is already in hand
L2 Harmonisation Group	Re-establish IHO/DGIWG Harmonisation Group
L3 Remap Models	Remap both existing models into new common model
L4 AML/TOD Products	Future releases of AML & TOD to make more extensive use of such a model
L5 Independent Encoding	The AML, TOD, ENC and DNC products should be described in an encoding independent manner using the UML modeling language to describe a content model for each product.
L6 Unique IDs	Object Level Unique IDs should be used to assist in updating and in product integration.
L7 Integrated Database	Producers should build integrated databases that can hold all of the objects defined in sets of products, independent of encodings – eg NIMA GIFD, UKHO V2F
L8 Repeatable Attributes	Full support for repeating attribute values
L9 Reference External Attributes	Full support for external references – text, pictures, multi-media files
S1 Feature Data Model	Do not restrict AML & TOD to ENC & DNC constructs
S2 Universal Unique IDs	Universal Unique IDs should be carried for all feature Objects in both ENC and DNC, and also for AML and TOD.
S3 Metadata	XML based auxiliary tables (ISO 19115 compliant) to carry further metadata to allow these products to be found in general product databases
S4 Feature Level Constructs	XML based auxiliary tables to carry aggregation and other feature level constructs
S5 Tighten the Testing Criteria	Geometry for interconversion; ENC – tighten Group 2 QC; DNC – tighten inter-coverage tests
S6 Geometric Inconsistencies	Correct DNC (and other products based on its VPF implementation) mode position problems Allow connected and isolated nodes to share position – it is standard VPF
DATA DICTIONARY	
L10 Interoperable Structures	Make the structures of the data dictionaries interoperable. Base on ISO TC211 structures and methods
L11 Table Adoption	Agree one to one alignments

L12 DGIWG Definition Adoption	DGIWG should adopt IHO definitions for hydrographic elements
L13 IHO Definition Adoption	IHO should adopt definitions for certain DGIWG elements
L14 Definition Revision	DGIWG and IHO should jointly revise and harmonise definitions
L15 DNC Alignment	Align DNC to modern FACC where advantageous
L16 ENC Alignment	IHO to consider changed definitions for ENC
L17 Feature Specialisation	Feature that are specialisation of other features need to be identified
L18 Collection Feature Objects	New collection feature objects should be included in FACC and S-57 to describe features composed of other features
S7 AML Data Dictionary	Establish AML data dictionary register; Under control of NATO AHHWG
S8 Extend "Code Space"	Extend IHO/DGIWG code spaces to facilitate assigning blocks of codes
L19 AML Register	Rationalise AML register with authoritative referencing to S-57 and FACC registers
L20 AML/TOD Relationship	Establish appropriate AML/TOD relationship
PRESENTATION	
L21 GeoSym Symbolisation	Make GeoSym fully S-52 compatible – symbols and rules
L22 GeoSym Rules	
L23 Neutral Portrayal Rules	Establish set of neutral portrayal rules based on ISO 19117 profile
METHODS	
L24 Neutral Methods	Define operations as part of features
L25 Free Text Data Fields	Change free text fields to enumerated data fields wherever possible
EXCHANGE OF DATA	
L26 Carry Neutral Data	Permit data that is neutrally described to be carried in either S-57 or VPF. That is, the existing encodings remain valid, and the content may be carried up to the limitations of the exchange format.
L27 XML Metadata	Add XML based metadata in accordance with ISO 19115 for data discovery
L28 XML UID	Add XML based data to permit 'unique ID' matching
L29 Auxiliary Collection Objects	Add XML based information to map collection objects between S-57 and FACC
L30 XML Encoding	Investigate XML as basis for alternate encoding

CONVERSION	
S9 Data Conversion	Only consider data conversion within a production environment
L31 Production Systems	Design production systems around common model, registers etc
L32 Mapping Schema	Develop standardised mapping schema from list of conversion rules
TESTING	
S10 “False Errors”	Develop a list of "false" errors that emanate from the DNC Validator or the ENC testing scheme corresponding to the relaxing of criteria or minor changes needed to improve conversion or production of two products from one source.
L33 Model Based Testing	Develop model based testing procedures
L34 Develop Test Data	Develop comprehensive test data