Harmonization of Nautical Information

Document No. D_WP4_3

Document Access: Public

Version: 1.0.1

Date: 20 December 2011

Author:

JEPPESEN
A BOEING COMPANY

Part-financed by the European Union (European Regional Development Fund and European Neighbourhood and Partnership Instrument)

DOCUMENT STATUS

Authors

Name	Organization
Raphael Malyankar	Jeppesen

Reviewers

Name	Organization
Cathrine Fierberg	Jeppesen
Attila Zabos	Jeppesen
Jens Schröder-Fürstenberg	BSH Hamburg, Germany
Karin Vaksdal	NHS
Odd-Aage Føre	NHS
Olav Haugen	NHS
Eivind Mong	Jeppesen

Approval of report

Name	Organization	Signature	Date

Document history

Revision	Date	Organization	Initials	Revised pages	Short description of changes
0.1	07-Nov-11	Jeppesen	RMM		First Draft
1.0	12-Dec-2011	Jeppesen	RMM	Various	Changes after draft review
1.0.1	20-Dec-2011	Jeppesen	RMM	Logos page and styles	Added logos, updated styles of title page & header/footer

Norwegian Hydrographic Service • Aalborg University, Denmark • Agency for Maritime and Coastal Services, Belgium • Danish Coastal Authority • Federal Maritime & Hydrographic Agency, Germany • Hjørring Municipality, Denmark • Jeppesen GmbH, Germany • Local Government, Denmark • Mälardalen University, Sweden • National Space Institute, Denmark • National Survey and Cadastre, Denmark • Natural Environment Research Council, United Kingdom • Norwegian Coastal Administration • Seazone Solutions Limited, United Kingdom •T-Kartor AB, Sweden • TU Delft, the Netherlands • UK Hydrographic Office

Table of Contents

Executive summary

This report summarizes research completed on the harmonization of nautical publications for the North Sea.

The harmonization project consists of converting selected information from official nautical publications published by national Hydrographic Offices into datasets conforming to the new S-100 framework for marine information, recently introduced by the International Hydrographic Organisation.

Information in nautical publications, being textual, narrative, unstructured and expressed in natural language, is difficult to express in object/attribute form. The IHO Standardisation of Nautical Publications Working Group (SNPWG) is working on a data model for nautical publications which is based on S-100. The Harmonisation phase converted unstructured textual information from nautical publications pertaining to one port each from Denmark, Germany, and Norway into structured data conforming to the IHO SNPWG data model, and produced XML datasets for the three ports. The project used the information requirements of a hypothetical digital mariner's routing guide to drive the selection of data from nautical publications. This phase of the project demonstrates that nautical publications information for disparate ports from the abovementioned North Sea nations can be harmonised by means of conversion into a single data model.

BLAST is an INTERREG IVB North Sea project involving 16 partners from 6 countries, including government organizations, universities and private companies. BLAST will demonstrate improvements in maritime safety, economy, and environmental management that can result from trans-national harmonization of land and sea data. There are 6 Work Packages in all. For more detail on BLAST, please refer to http://www.blast-project.eu/. The work described in this document is a part of Work Package 4 of the BLAST project.

1. Preface

1.1. The BLAST (Bringing Land and Sea Together) project

BLAST is an Interreg IVB North Sea project promoting maritime safety in the North Sea region. It involves 16 partners from 6 countries, including government organizations, universities and private companies. The BLAST partners are collaborating to demonstrate the potential for improvements in maritime safety, economy, and environmental management that can result from trans-national harmonization and integration of land and sea data. BLAST is divided into 6 Work Packages and a variety of projects from late 2009 through 2012. The BLAST Website summarizes BLAST and gives detail on its work packages - http://www.blast-project.eu/.

1.2. BLAST Work Package 4: Navigating the North Sea

The projects in WP 4, conducted by Jeppesen GmbH and partners, examine the collection, processing, and publishing of nautical information. The Jeppesen projects started with primary research to identify the current state of the art in Denmark, Germany, and Norway. This enabled the partners to identify what works well today and where there trans-nationally harmonized data could make it easier for the North Sea community to maintain a clear, current view of maritime conditions. The third project, harmonization of Nautical Information, is the focus of this paper.

1.2.1. Project 1: Survey Current State of the Art

The report summarized research into the current state of the art in nautical information management systems in three North Sea ports. The research is related to three projects in BLAST Work Package 4 conducted by Jeppesen GmbH, Mälardalen University, the Hydrographic Offices of Norway, Denmark, and Germany, and the Coastal and Port Authorities in the three subject ports: Hirtshals, DK, Stavanger, NO, and Wilhelmshaven, DE.

1.2.2. Project 2: Maritime Data Collection System Web Demonstrator

The Maritime Data Collection System (MDCS) project demonstrated a web-based solution prototype of trans-nationally harmonized maritime data collection, validation, and publications. MDCS requirements definition started in Phase 1.

1.2.3. Project 3: Harmonization of Nautical Information

The Harmonization of Nautical Information project demonstrated conversion of selected nautical information for the three test ports into XML datasets conforming to a common data model. The information consists of extracts from publications produced by the Hydrographic Offices and Coastal Authorities in Norway, Germany, and Denmark. The common data model conforms to the new IHO S-100 standard for hydrographic information. North Sea Hydrographic Offices have a central role in the definition of the standard. This project is a practical test of applying the IHO work to nautical publications information.

1.2.4. Project 4: Digital Mariners Routing Guide Web Demonstrator

The Digital Mariners Routing Guide project aims to demonstrate the potential benefits of using trans-nationally harmonized data in products that can improve maritime efficiency and safety. It uses the datasets produced in the earlier phases. Thus the North Sea DMRG builds upon what was learned in the prior projects. The 2008 Baltic Sea Digital Mariners' Routing Guide and existing North Sea Routing Guides will be referenced. This project will start in late 2011 and complete in early 2012.

2. Introduction

2.1. Background

The modern mariner uses navigation information from several sources. Two important sources are nautical charts (paper or electronic) and text publications such as sailing directions and port guides. Charts provide depictions of geographic data. Nautical publications provide general guidance and non-geographic information, such as extracts from navigation and other regulations, information about vessel traffic services, radio services, pilot services and pilot requirements, natural conditions, hazards and obstructions, overviews of coastal areas and routes, and special notes and supplemental information about geographic features. Nautical publications also include emphases and rearrangements of information which is available in charts.

In addition to charts and nautical publications, mariners use recent or real-time information including weather, navigational warnings, etc. This may be obtained by one form or another of radio communications, and more recently also over the Internet.

Obviously the Internet can be used for disseminating many types of information, including recent or (near) real-time information, digital charts, publications, and updates to them, and it is in fact being so used by various producers including hydrographic offices, coastal authorities, ports, government agencies, and private sources. Traditionally, nautical publications were distributed in printed form. Digital formats have been introduced relatively recently for charts (ENCs) and are being increasingly used. Digital formats for nautical publications have generally been limited to digitized forms (e.g., PDF files or HTML Web pages) of the corresponding paper documents or database-backed web presentations (e.g., lists of chart corrections or inoperable navigation aids).

A data model for hydrographic data was introduced by the IHO in the 1990's in the S-57 standard [S57], and is currently used for ENCs. The IHO has since developed a newer framework standard "S-100: Universal Hydrographic Data Model" [S100], which is intended to support other types of navigation information and other data products in addition to ENCs. S-100 is a framework standard from which standards for different data types and data products are derived. Examples of such derived standards are the S-101 and S-102 Product Specifications, both currently under development, which describe data models for ENC and high-density bathymetry data respectively.

Chart data and bathymetry are well-structured data which can easily be expressed using object-attribute data models, but information in nautical publications, being textual, narrative, unstructured and expressed in natural language, is difficult to express in object/attribute form. The IHO Standardisation of Nautical Publications Working Group (SNPWG) is working on a data model for nautical publications which is based on S-100. The Harmonisation phase converted unstructured textual information from nautical publications pertaining to one port each from Denmark, Germany, and Norway into structured data conforming to the IHO SNPWG data model, and produced XML datasets for the three ports. The project used the information requirements of a hypothetical digital mariner's routing guide to drive the selection of data from nautical publications. This phase of the project demonstrates that nautical publications information for the 3 participating North Sea nations and for ports of different size can be harmonised by means of conversion into a single data model.

2.2. Acronyms and Abbreviations

BSH	Bundesamt für Seeschifffahrt und Hydrographie
DaMSA	Danish Maritime Safety Administration (see also FRV)
ENC	Electronic Navigation Chart
FRV	Farvandsvæsenet (see also DaMSA)
IHO	International Hydrographic Organisation
IMO	International Maritime Organisation
KMS	Kort & Matrikelstyrelsen (National Survey and Cadastre)
MDCS	Maritime Data Collection System
NCA	Norwegian Coastal Administration (Kystverket)
NHS	Norwegian Hydrographic Service (Statens Kartverk Sjø)
SNPWG	Standardisation of Nautical Publications Working Group (IHO)
TSMAD	Transfer Standard Maintenance and Development Working Group (IHO)
VTS	Vessel Traffic Service
WSA	Water and Shipping Agency (Wasser- und Schifffahrtsamt)
WSD	Waterway and Shipping Directorate (Wasser- und Schifffahrtsdirektion)
WSV	Wasser- und Schifffahrtsverwaltung des Bundes

2.3. References

[BLAST1] State of the art report: Nautical information management. BLAST Deliverable D_WP4_1, 2010. URL: http://www.blast-project.eu/media.php?file=283

[DRGPS] BLAST Digital Routing Guide Product Specification. (Under development.)

[S49] Standardization of Mariner's Routeing Guides. Special Publication No. S-49, Edition 2.0. April 2010. International Hydrographic Bureau, Monaco.

[S57] International Hydrographic Organization. S-57: IHO Transfer Standard for Digital Hydrographic Data, ed. 3.1. November 2000. (latest supplement June 2009)

[S100] International Hydrographic Organization. S-100: Universal Hydrographic Data Model, Edition 1.0.0. International Hydrographic Bureau, Monaco, 2010.

[SW] Wiki maintained by Capt. Schröder-Fürstenberg for Standardization of Nautical Publications Working Group discussions, URL: http://www.fuerstenberg-dhg.de/mediawki/index.php

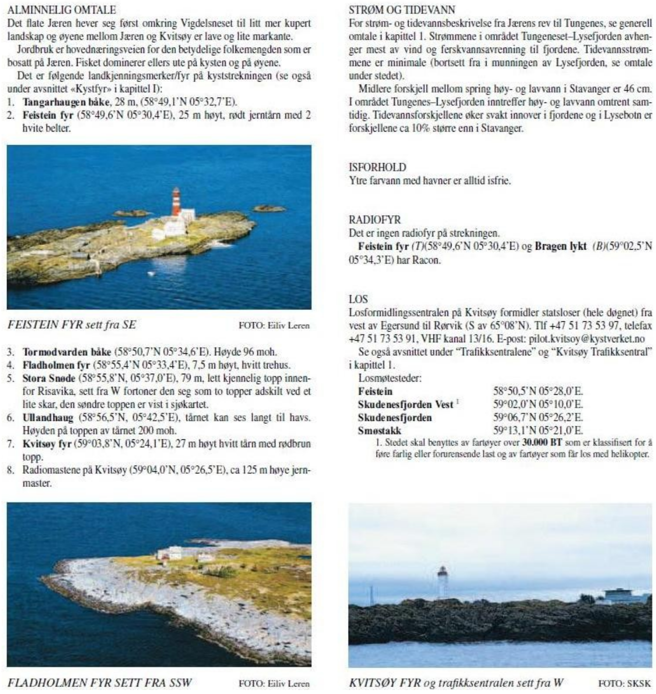
[WG2011] IHO S-100: The universal Hydrographic data Model. R. Ward and B. Greenslade, International Hydrographic Organisation. IHO Information paper, 2011. URL: http://www.iho-ohi.net/iho_pubs/standard/S-100/S-100_Info.htm

3. Source Material

The selection of source material was guided by the eventual application, namely the digital routeing guide to be developed in the next phase of BLAST. Source material consisted mainly of sailing directions or the equivalent, and electronic nautical charts. The ENC files were used mainly to determine the geographic locations with which information from the publications is associated. The primary nautical publications used were:

1. Den Norske Los, Volume 3: Jærens rev – Stad, 2006, Statens Kartverk; chapter 2. This is the sailing directions for the Norwegian coast in the vicinity of Stavanger.
2. Den Norske Los, Volume 1: Alminnelige opplysninger, 2010, Statens Kartverk, Norway, chapters 2 & 3. Contains general information and regulations for Norwegian waters.
3. Nordsee-Handbuch, Südöstlicher Teil, Bundesamt für Seeschifffahrt und Hydrographie, Part C, chapters 4 and 5. These chapters comprise BSH sailing directions for the German North Sea coast in the vicinity of Wilhelmshaven and the Jade basin.
4. Nordsee-Handbuch, Südöstlicher Teil, Bundesamt für Seeschifffahrt und Hydrographie, Part A. This contains general information and navigation regulations for German waters.
5. Vessel Traffic Services (VTS) Guide, Germany, 2010, Bundesamt für Seeschifffahrt und Hydrographie. Information about vessel traffic services in the vicinity of Wilhelmshaven.
6. Revierfunkdienst Nordsee und Englischer Kanal, Bundesamt für Seeschifffahrt und Hydrographie, 2011. List of Radio Stations for the North Sea and English Channel. This was used for information about radio stations broadcasting marine information for the German North Sea coast.
7. Nautischer Funkdienst 2010/11: This was used for information about radio broadcasting marine information for the German part of the German Bight.
8. Danske Havnelods, web page for the port of Hirtshals, published only on the web by KMS. This web page provides information about facilities, services, conditions, etc., in the port of Hirtshals and the immediate vicinity.
9. Den Danske Lods – II, Kort & Matrikelstyrelsen. This provides general information and regulations for Danish waters.

Structure and volume of the source texts varied widely in their descriptions of the port areas and approaches. The Norwegian sailing directions for the Stavanger area (in *Den Norske Los*, Volume 3) are written as plain text in a narrative style, include several lists (of landmarks, pilot boarding places, etc.) and describe the approaches from seaward as well as providing details about berth sizes and port services. A sample page is provided in Figure 1. KMS has replaced printed sailing directions for port areas with web pages for individual ports, available on the KMS web site as *Danske Havnelods*; the page for Hirtshals is shown in Figure 2. The structure is less of a flowed narrative compared to Den Norske Los, and is distinguished from the other two in using pictograms to convey some of its information about port services. BSH provides the equivalent information in Part C of the *Nordsee Handbuch* – a sample page is shown in Figure 3. This is structured as short sentences and sentence fragments which contain the essential information, and makes intensive use of visual style cues such as fonts and indentation to arrange associated information elements together and indicate the areas and vessels to which specific rules apply. Den Norske Los tends to provide area-specific regulations in full, compared to the other two; the Nordsee Handbuch presents them as text fragments prescribing or forbidding actions, under heads indicating the conditions under which a the specific prescription/prohibition applies. All three arrange the information under similar heads and subheads.



ALMINNELIG OMTALE

Det flate Jæren hever seg først omkring Vigdelsneset til litt mer kupert landskap og øyene mellom Jæren og Kvitsøy er lave og lite markante.

Jordbruk er hovednæringsveien for den betydelige folkemengden som er bosatt på Jæren. Fisket dominerer ellers ute på kysten og på øyene.

Det er følgende landkjenningsmerker/fyr på kyststrekningen (se også under avsnittet «Kystfyr» i kapittel I):

1. **Tangarhaugen båke**, 28 m, (58°49,1'N 05°32,7'E).
2. **Feistein fyr** (58°49,6'N 05°30,4'E), 25 m høyt, rødt jerntårn med 2 hvite belter.

FEISTEIN FYR sett fra SE FOTO: Eiliv Leren

3. **Tormodvarden båke** (58°50,7'N 05°34,6'E). Høyde 96 moh.
4. **Fladholmen fyr** (58°55,4'N 05°33,4'E), 7,5 m høyt, hvitt trehus.
5. **Stora Snøde** (58°55,8'N, 05°37,0'E), 79 m, lett kjennelig topp innenfor Risavika, sett fra W fortoner den seg som to topper adskilt ved et lite skar, den søndre toppen er vist i sjøkartet.
6. **Ullandhaug** (58°56,5'N, 05°42,5'E), tårnet kan ses langt til havs. Høyden på toppen av tårnet 200 moh.
7. **Kvitsøy fyr** (59°03,8'N, 05°24,1'E), 27 m høyt hvitt tårn med rødbrun topp.
8. Radiomastene på Kvitsøy (59°04,0'N, 05°26,5'E), ca 125 m høye jernmaster.

FLADHOLMEN FYR SETT FRA SSW FOTO: Eiliv Leren

STRØM OG TIDEVANN

For strøm- og tidevannsbeskrivelse fra Jærens rev til Tungenes, se generell omtale i kapittel 1. Strømmene i området Tungeneset–Lysefjorden avhenger mest av vind og ferskvannsavrenning til fjordene. Tidevannsstrømmene er minimale (bortsett fra i munningen av Lysefjorden, se omtale under stedet).

Midlere forskjell mellom spring høy- og lavvann i Stavanger er 46 cm. I området Tungenes–Lysefjorden inntreffer høy- og lavvann omtrent samtidig. Tidevannsforskjellene øker svakt innover i fjordene og i Lysebotn er forskjellene ca 10% større enn i Stavanger.

ISFORHOLD

Ytre farvann med havner er alltid isfrie.

RADIOFYR

Det er ingen radiofyr på strekningen.

Feistein fyr *(T)*(58°49,6'N 05°30,4'E) og **Bragen lykt** *(B)*(59°02,5'N 05°34,3'E) har Racon.

LOS

Losformidlingssentralen på Kvitsøy formidler statsloser (hele døgnet) fra vest av Egersund til Rørvik (S av 65°08'N). Tlf +47 51 73 53 97, telefax +47 51 73 53 91, VHF kanal 13/16. E-post: pilot.kvitsoy@kystverket.no

Se også avsnittet under "Trafikksentralene" og "Kvitsøy Trafikksentral" i kapittel 1.

Losmøtesteder:

Feistein	58°50,5'N 05°28,0'E.
Skudenesfjorden Vest [1]	59°02,0'N 05°10,0'E.
Skudenesfjorden	59°06,7'N 05°26,2'E.
Smøstakk	59°13,1'N 05°21,0'E.

1. Stedet skal benyttes av fartøyer over **30.000 BT** som er klassifisert for å føre farlig eller forurensende last og av fartøyer som får los med helikopter.

KVITSØY FYR og trafikksentralen sett fra W FOTO: SKSK

Figure 1. Sample page from Den Norske Los

Hirtshals Havn

Sidste opdateringer • Tekst: 29-9-2010 - Plan 1: 27-2-2008

Beliggenhed
Skagerrak
57°35,7'N 9°57,7'E - kort 92

fotograferet august 2004

Havnen
ejes af Hjørring Kommune.
Havnen består af følgende bassiner: Forhavnsbassin, Vestbassin, Mellembassin I, Mellembassin II, Mellembassin III, Østbassin I og Østbassin II med flydedok.
I Forhavnsbassin findes en bådehavn fortrinsvis for hjemmehørende lystfartøjer og fartøjer for bierhvervsfiskere. Bådehavnen administreres af Hirtshals Bådeklub.
I Vestbassin findes bådebro med plads til ikke-hjemmehørende lystfartøjer. Bådebroen administreres af foreningen Hirtshals Marina.

Dybder
Se plan.
Som følge af sandvandring kan de opgivne dybder ikke altid påregnes at være til stede foran havnen, i Forhavnsbassin eller overalt i havnebassinerne.

Figure 2. Web page for Hirtshals from Danske Havnelods

C 2 German Bight and Helgoland

C 2.1 German Bight and Helgoland

C 2.1.1 German Bight

Approach

Routeing Systems
IMO-Tanker route see chapter A 2.1
TSSs „German Bight Western Approach„, „Terschelling–German Bight„ and „Jade Approach„
navigate according to Rule 10 of the COLREGS

Route planning see also chapter A 2.1
refer „VTS Guide Germany„ (BSH-Nr. 2011)
danger of crossing traffic at the junction of the TSSs Terschelling-German Bight and Jade Approach
inshore traffic zone
navigation in several parts affected by national regulation
prohibited for transit traffic which can navigate safely in the appropriate TSS areas

Landmarks see also C 2.1.2 and figures

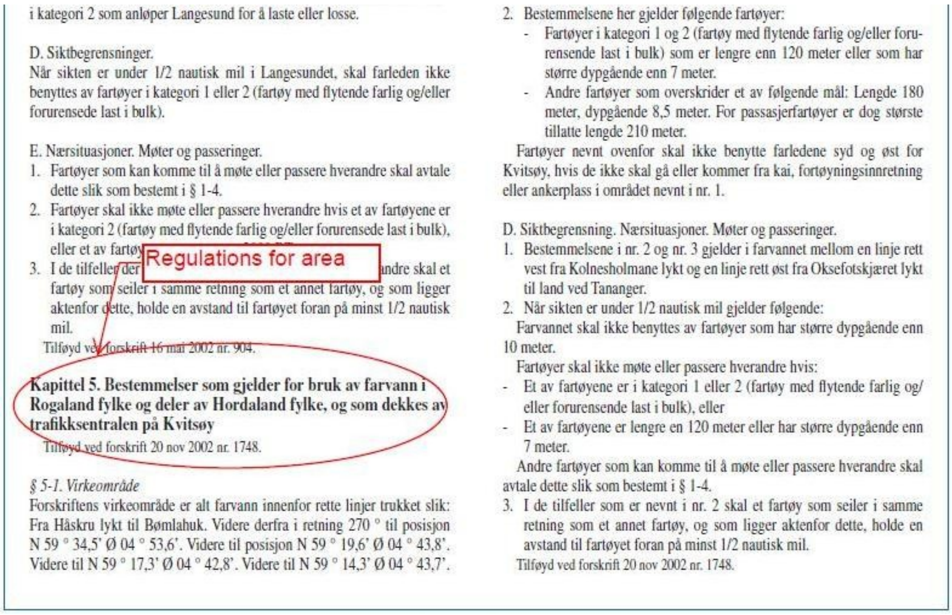
Unmanned light vessel GB
Figure C 2.1.1

Shipping restrictions see also chapter A 1.3

Military Exercise Area see also chapter A 1.6
submarine exercise area
NW of Helgoland
Area Weser

Figure 3. Sample page from Nordsee Handbuch, Part C

All three hydrographic offices provide general information and the text of regulations in volumes or chapters different from the area-specific information – Volume 1 of *Den Norske Los* for Norway, a printed publication called *Den Danske Lods* for Denmark (not currently available on the Web), and Part A of the *Nordsee Handbuch* for Germany. The presentation of this information is more consistent across the three countries, is more of a narrative text format. A sample page is given in Figure 4.



i kategori 2 som anløper Langesund for å laste eller losse.

D. Siktbegrensninger.
Når sikten er under 1/2 nautisk mil i Langesundet, skal farleden ikke benyttes av fartøyer i kategori 1 eller 2 (fartøy med flytende farlig og/eller forurensede last i bulk).

E. Nærsituasjoner. Møter og passeringer.
1. Fartøyer som kan komme til å møte eller passere hverandre skal avtale dette slik som bestemt i § 1-4.
2. Fartøyer skal ikke møte eller passere hverandre hvis et av fartøyene er i kategori 2 (fartøy med flytende farlig og/eller forurensede last i bulk), eller et av fartøy
3. I de tilfeller der ... andre skal et fartøy som seiler i samme retning som et annet fartøy, og som ligger aktenfor dette, holde en avstand til fartøyet foran på minst 1/2 nautisk mil.

Tilføyd ved forskrift 16 mai 2002 nr. 904.

Kapittel 5. Bestemmelser som gjelder for bruk av farvann i Rogaland fylke og deler av Hordaland fylke, og som dekkes av trafikksentralen på Kvitsøy

Tilføyd ved forskrift 20 nov 2002 nr. 1748.

§ 5-1. Virkeområde
Forskriftens virkeområde er alt farvann innenfor rette linjer trukket slik: Fra Håskru lykt til Bømlahuk. Videre derfra i retning 270 ° til posisjon N 59 ° 34,5' Ø 04 ° 53,6'. Videre til posisjon N 59 ° 19,6' Ø 04 ° 43,8'. Videre til N 59 ° 17,3' Ø 04 ° 42,8'. Videre til N 59 ° 14,3' Ø 04 ° 43,7'.

2. Bestemmelsene her gjelder følgende fartøyer:
 - Fartøyer i kategori 1 og 2 (fartøy med flytende farlig og/eller forurensende last i bulk) som er lengre enn 120 meter eller som har større dypgående enn 7 meter.
 - Andre fartøyer som overskrider et av følgende mål: Lengde 180 meter, dypgående 8,5 meter. For passasjerfartøyer er dog største tillatte lengde 210 meter.

Fartøyer nevnt ovenfor skal ikke benytte farledene syd og øst for Kvitsøy, hvis de ikke skal gå eller kommer fra kai, fortøyningsinnretning eller ankerplass i området nevnt i nr. 1.

D. Siktbegrensning. Nærsituasjoner. Møter og passeringer.
1. Bestemmelsene i nr. 2 og nr. 3 gjelder i farvannet mellom en linje rett vest fra Kolnesholmane lykt og en linje rett øst fra Oksefotskjæret lykt til land ved Tananger.
2. Når sikten er under 1/2 nautisk mil gjelder følgende:

Farvannet skal ikke benyttes av fartøyer som har større dypgående enn 10 meter.

Fartøyer skal ikke møte eller passere hverandre hvis:
- Et av fartøyene er i kategori 1 eller 2 (fartøy med flytende farlig og/eller forurensende last i bulk), eller
- Et av fartøyene er lengre en 120 meter eller har større dypgående enn 7 meter.

Andre fartøyer som kan komme til å møte eller passere hverandre skal avtale dette slik som bestemt i § 1-4.

3. I de tilfeller som er nevnt i nr. 2 skal et fartøy som seiler i samme retning som et annet fartøy, og som ligger aktenfor dette, holde en avstand til fartøyet foran på minst 1/2 nautisk mil.

Tilføyd ved forskrift 20 nov 2002 nr. 1748.

Figure 4. Sample regulations from Den Norske Los

4. Data Model

4.1. Modelling Framework

The “Universal Hydrographic Data Model” described in S-100 is a framework intended to support data modelling for a variety of data sources, products, and applications. The model is based on ISO/TC211 standards for geographic information. The main characteristics of the model are objects, attributes, and relationships, described below.

A **feature** is an abstraction of real world phenomena, e.g., charted features such as a restricted area, harbour area, navigation aid, etc. In the S-100 model, *GF_FeatureType* is a meta-class that is instantiated as classes that represent individual feature types. A certain feature type is the class used for all instances of that feature type. The instances of a class that represents an individual feature type are called feature instances. Feature types are the equivalent of classes and feature instances are equivalent to objects in UML.

An **information type** is an identifiable object type that can be associated with features and information objects in order to carry information pertaining to the associated features, or qualify the information in the other information object. A regulation pertaining to passage requirements for a fairway is an example of an information type which is associated with the geographic feature type representing the fairway. *S100_GF_InformationType* is the class intended to be the metaclass for information types in S-100.

An **attribute** is a property of a feature or information type. **Simple Attributes** can be enumerations, code lists, simple types (e.g. integer or character string), or special types (dates and times).

Complex attributes are properties of a feature which can be divided into multiple sub-attributes and are used where objects have properties that better fit a hierarchical structure. They provide a better construct than simple attributes for encoding list attributes on objects, such as light sectors.

Aggregations are groups of feature instances and are used represent features that are related to each other.

A **feature association** is a relationship that links instances of one feature type with instances of the same or a different feature type. Each relationship has a name and two roles thus giving a more informative representation of the real-world relationships in the domain being modelled.

An **information association** is a relationship that links instances of a feature or information type with an information type. Each relationship has a name and role in order to provide a more informative representation of relationships.

The S-100 specification provides for a **feature concept dictionary** (FCD), a collection of descriptions of the objects and attributes in a defined application area or domain (e.g., nautical charts). In the FCD, attributes are not bound to objects; this binding is done separately for each application area data product, in the **feature catalogue** of a **product specification** (section 5). This allows reuse of objects and attributes in different data products in the same or other domains.

More information about the framework is available from IHO publications [S100, WG2011].

4.2. Nautical Publications Information Data Model

The IHO working group on Standardization of Nautical Publications (SNPWG) has developed an information model based on the S-100 framework for the type of information found in textual nautical publications (as opposed to nautical charts). This model is currently documented online at the SNPWG Wiki [SW]. It currently defines 22 new geographic features, extends 7 geographic features defined in the current hydrographic data model [S-57], and defines 10 new information types (a concept introduced in S-100). It includes 148 simple attributes (from S-57 and newly defined) and defines 17 complex attributes (another concept introduced in S-100).

The scope of this model is far larger than digital mariners routing guides. The geographic objects range from radio stations to port security areas. The information objects range from regulations to address/contact information, ship reports, hours of work, and specification of classes of vessels to which a rule applies.

5. Product Specification

A “product specification” in the context of the S-100 framework is a specification of a data product for a specific application area. An S-100 product specification consists of a description of the objects and relationships used in modelling a specified domain, accompanied by specifications of the structure, metadata, and encapsulation of data sets in an application area, and rules for portrayal, data capture and encoding, data quality, and product delivery. An important part of the product specification is the **feature catalogue**, which specifies the attributes of each object in the domain and the allowed ranges of attribute values.

The scope of the product specification was determined by the next planned BLAST activity: development of a digital mariner's routing guide. Features and attributes were selected for inclusion in the feature catalogue based on the guidelines for mariners' routing guides provided by the IHO publication "Standardization of Mariner's Routeing Guides" [S49], augmented by selected features from the SNPWG model which were are needed to capture the types of information commonly available in routing guides currently published by hydrographic offices (BSH and UKHO). The feature catalogue contains 95 features and 40 attributes from the ENC domain, and 18 features, 109 attributes, and 9 information objects from the nautical publications domain.

The S-100 standard allows for different formats for data sets. The encoding for this product specification was defined using an XML application schema derived from GML 3.2.1. The encoding is capable of coding geographic features, information objects, simple and complex attributes, feature associations, information associations. Figure 5 shows the XML encoding of the feature type for Marine Service – the components shown are a pointer to the object geometry (XML element *spatialref*), association to information objects (element *informationAssociation* in the figure), certain simple attributes common to all feature types (in the boxes for *AttributeGroupB* and *AttributeGroupC*), and attributes specific to this feature (*categoryOfMarineService*, *serviceAccessProcedure*, etc.). The types (*MarineServiceType*, *CATMSVType, S100Feature:AbstractS100FeatureType*, etc.) used in this definition are defined elsewhere in the schema or in accompanying schema files.

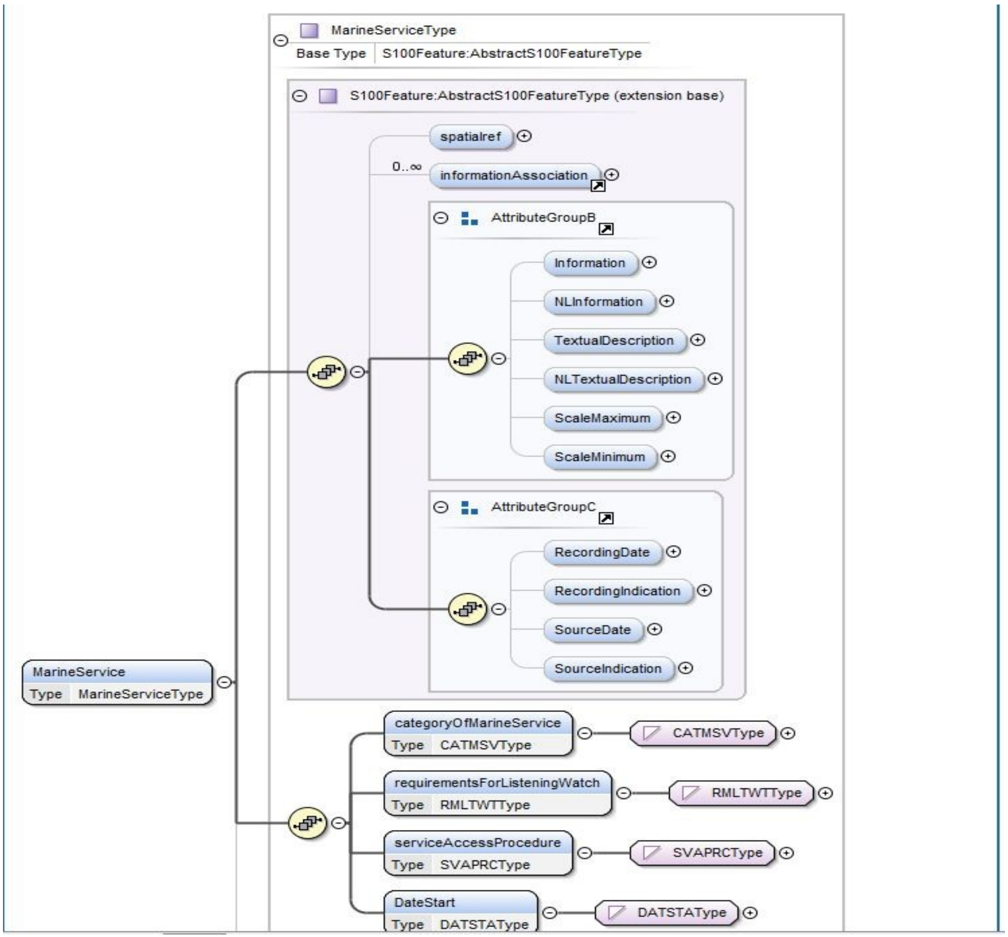


Figure 5. Typical feature encapsulation in XML schema

Figure 6 below shows the XML code corresponding to a VTS area. It uses the *MarineService* object of Figure 5 and includes, in order, the coordinates of the polygon (in decimal degrees with a multiplication factor of 10^7 applied), associations to three information objects (2 regulations objects and 1 for contact details), and simple attributes describing the type of feature (value 1 = VTS), requirements for listening watch for this area, and the effective date of the VTS area.

```
- <dmrg:MarineService gml:id="mrnsrv01">
  - <spatialref>
    - <gml:Polygon gml:id="a01">
      - <gml:exterior>
        - <gml:LinearRing>
            <gml:posList srsDimension="2">052398333 595491667 051848333 595748333 048933333
              595750000 047300000 593266667 047133333 592883333 047283333 592383333
              050900000 590933333 052483333 589816667 053600000 587500000 054933333
              587500000 052398333 595491667</gml:posList>
          </gml:LinearRing>
        </gml:exterior>
      </gml:Polygon>
    </spatialref>
    <S100Feature:informationAssociation gml:id="mrnsrv01andreglts01A" xlink:href="#reglts01A" />
    <S100Feature:informationAssociation gml:id="mrnsrv01andreglts01B" xlink:href="#reglts01B" />
    <S100Feature:informationAssociation gml:id="mrnsrv01andcondet02" xlink:href="#condet02" />
    <dmrg:categoryOfMarineService>1</dmrg:categoryOfMarineService>
    <dmrg:requirementsForListeningWatch>Vessel in motion or at anchor in the area shall keep a continuous listening
      watch on the traffic unit's working frequency. Permission to change to another frequency or to cancel the watch
      must be obtained from "KVITSØY CONTROL".</dmrg:requirementsForListeningWatch>
    <dmrg:serviceAccessProcedure>VHF channel 18 (S), VHF channel 19 (N)</dmrg:serviceAccessProcedure>
    <dmrg:DateStart>2003-01-01T20:00:00+01:00</dmrg:DateStart>
  </dmrg:MarineService>
```


Figure 6. XML code for typical feature

More information is available in the product specification document [DRGPS].

6. Information Mapping and Conversion

6.1. Information mapping process and results

Information mapping is, for the purposes of this project, the process of identifying the objects and defining their attributes from the source material identified earlier. Examples of information mapping are shown in Figure 7, which demonstrates mappings from the port sections of the BSH *Nordsee Handbuch* and NHS *Den Norske Los* to objects for Regulations, Contact Details, Shoreline Construction, and Supplies.

The information mapping work for this project was performed by project participants at BSH, NHS, and KMS, and reviewed by Jeppesen. The three hydrographic offices determined which portions of the material listed in section 3 were relevant to the routing guide. They then prepared Excel spreadsheets conforming to a template determined by Jeppesen and the hydrographic offices, listing the objects, attributes and their values, and indicating associations between objects. Coordinates for geographic features were determined from the text if available; if not, the coordinates were obtained from the appropriate ENC or created in an ENC editing tool. The results were reviewed by Jeppesen and several minor corrections were made, which were in turn reviewed by the hydrographic offices. The result was spreadsheets containing the information mappings in tabular form, which were converted into XML representations of the mapped information.

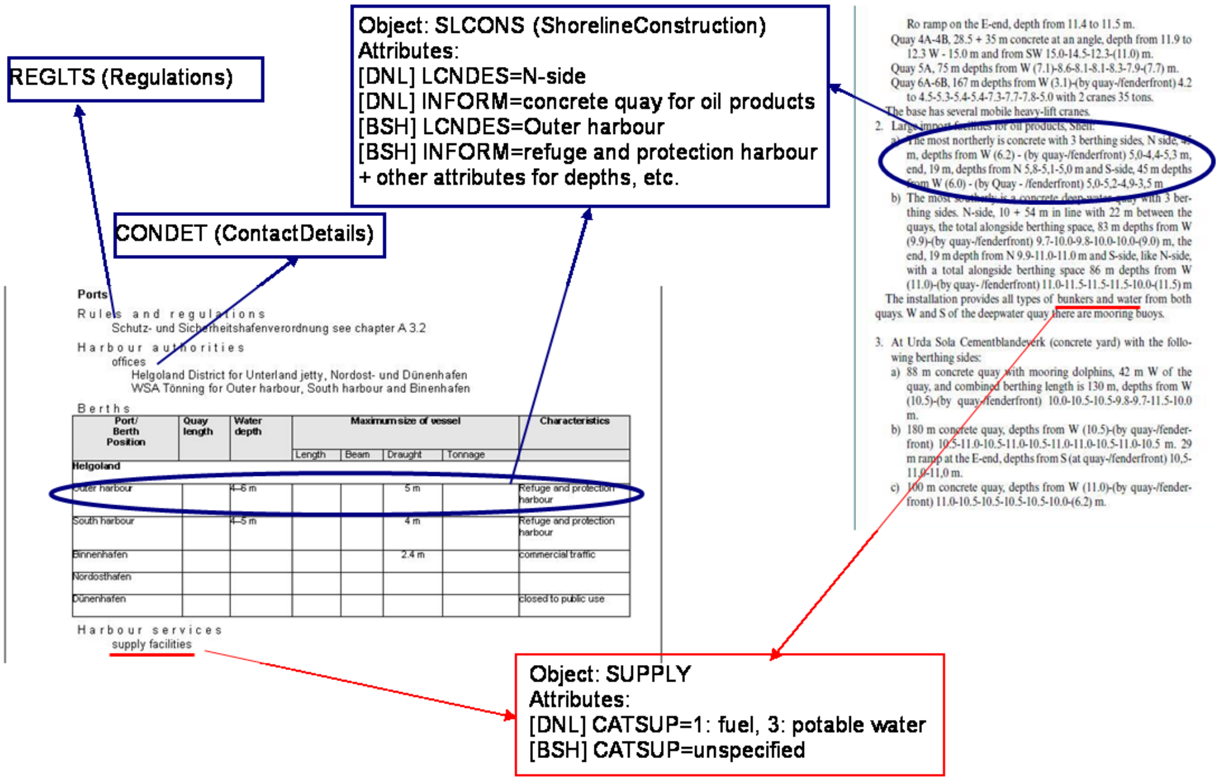


Figure 7. Examples of text mapping

6.2. Information Mapping Experiences

HO partners reported their experiences with mappings. The main points reported were:

Overall Effort: Information mapping took significant effort, approximately 73 hours for the BSH portion of the work (which produced approximately 75% of the total end result data product).

Effort needed for spatial linkages: A large fraction of the time was spent on linkages to geometry, in the form of geographic features or spatial objects. BSH reported 18 hours for determining ENC feature object identifiers and 11 hours for assigning or creating coordinates for geographic features. The other two HOs took relatively lower proportions of effort, since they had proportionately fewer geographic features and in the case of NHS preferred to approximate complex geometry by a rectangle. Such approximate geometry was created where the existing geometry was considered too complex or not discoverable in the ENC. The unexpectedly high level of effort was attributed by the mappers to the fact that it is difficult to create precise and correct geometry, and also to find the correct geometry in the ENC. Mappers noted that sometimes a single geometry might not always apply to all vessels. One suggestion was to use feature object identifiers, which are supposed to uniquely identify a feature. It was noted that not all of the HOs maintain the feature object identifiers.

Place location: Nautical publications, in particular sailing directions, often identify locations verbally instead of by coordinate, e.g., by naming a buoy instead of providing its coordinates. It is often difficult to locate such named points in the ENCs; if the name is not coded in the ENC it is not possible to locate it without external (local or personal) knowledge about the area.

Geographic features not in chart: There are mismatches between the ENC and the nautical publications, in the form of locations or features mentioned in the nautical publications (by name or otherwise) not being in the ENC. It was not always clear which source was the correct one.

Chart discrepancies: The mapping process led to the discovery of discrepancies in the charts, for example a feature displaced from its actual location.

Proliferation of spatial objects: Different clauses in a regulation may apply to different sections of a waterway. The question arises whether it is better to define distinct spatial objects for sections where specialised rules apply, or to attach a larger chunk of text to fewer, larger spatial objects. In such situations a choice must be made between an overly complex result which requires more labour and an overly simple result which conflates information items which are distinct in the real world and may therefore be less useful to the mariner.

Scope of routing guide: The different HOs chose to map to different depth of detail. NHS chose to define the work area from fairways in open sea to the pilot station, not all the way to the fender line. BSH and KMS chose to include this area as well as selected port facilities. Traditional routing guides do not include details about specific port facilities, berths, etc., It remains to be decided whether a digital routing guide, which is able to hide and show portions of data at will, allow zooming and panning, and implement a multi-page routing guide can usefully include such information. This proposition will be tested during development of the digital routing guide in the next phase of BLAST.

Extracts of regulations: Preparing extracts of regulations and laws was reported as a special point of concern for NHS and KMS. Sailing directions tend to contain regulations expressed in relatively long paragraphs, perhaps reproduced from the official legal publication. Mapping lengthy text regulations to an object/attribute model introduces both the desire and need to create shorter texts, perhaps as sentence fragments, and often requires preparing an extract or drafting a short paragraph or sentence based on the original text. Sailing directions for the German portion of the total avoided such problems largely because the process of extraction and condensation had long been done by BSH in the course of converting their sailing directions to the new BSH format, described in section 3. In the case of Norway, some information in laws or regulation (usually about exceptions or applicability) was omitted because it did not fit the data model. This means that laws and regulations are not comprehensively encoded and the portions which are in the data set are simplified versions of the source material. Further, sometimes the original text itself leaves room for interpretation, making it hard to encode as discrete attribute values. One solution mentioned was to include the text or laws and regulations in full in the exchange set as a reference.

Translations: Since the DMRG is to be created in English there was also the need for translation and there are also concerns about whether the meaning of the original was preserved in translation. HO partners noted that the conversion process might end up changing such a text in splitting it up or translating it to another language.

Capturing rule-like information: All three HO partners reported difficulties capturing regulations, etc., in the object/attribute model due to the absence of model elements which can express the precise conditions under which a regulation applies. Regulations are often written to fit specific circumstances which were not considered general enough to warrant capture in the SNPWG model. This usually happens in the form of using special terms or vessel characteristics to specify vessels to which a regulation applies. There are a very large number of variations possible, and SNPWG modelers decided to limit the complexity of the model by not trying to cover every possible characteristic value and nuance. Instead, there is provision for a catchall text statement of special characteristics or conditions, which can be used. Apart from this issue, the S-100 framework currently lacks convenient modeling elements for expressing procedural and rule-based knowledge (i.e., writing “if/then” statements). The object-attribute model must be used for capturing information more naturally expressed as if/then rules. It is theoretically possible to do this, but the non-intuitive structures required greatly increase the difficulty of information mapping.

Extensions to the data model: A few extensions or adaptations to the SNPWG data model were needed, such as modeling of graphic images included in nautical publications. These were duly developed and passed on to SNPWG for consideration.

Tools: Better GIS and product editing tool support for the S-100 model was needed, for example functionality to automatically assign identifiers when a new object is defined. Tools specialized for S-100 are not currently available due to the newness of the S-100 data framework.

6.3. Transfer sets

The mapped and cleaned data was converted into data sets conforming to the encoding specification described in Section 5, using the spreadsheet’s string manipulation functions. The results were transferred to a commercial XML editor and corrected where needed to turn them into valid XML. The end result was an XML data set file for each of the HO partners. This was in turn reviewed by the HO partner and corrected if necessary. The results can be viewed in an XML editor tool as well as an off-the-shelf GML viewer. Typical XML code for a feature is shown in Figure 6 above. The figures below show the area features[1] for each data set and the corresponding data for a selected feature (analogous to a pick report in an ENC display). Certain features with very large polygons (NAVTEX station areas, radio service areas, weather forecast areas) were excluded from the BSH display to allow smaller areas to appear. The BSH data set has over thirty feature types, and several other types are too small to be visible. Overlapping polygons result in several other feature types being overlaid and not visible in Figure 9 and to a smaller extent in Figure 10. Point features such as landmarks are not shown.

[1] Colours: Sea Area: lightest blue; Pilot Service area: medium blue; Marine Service: dark blue; Military Practice area: dark pink; Restricted Area: light pink; Caution Area; red; Port Area: yellow; Natural Conditions area: green. The purple area in Figure 9 is a highlight of the picked Marine Service area object.

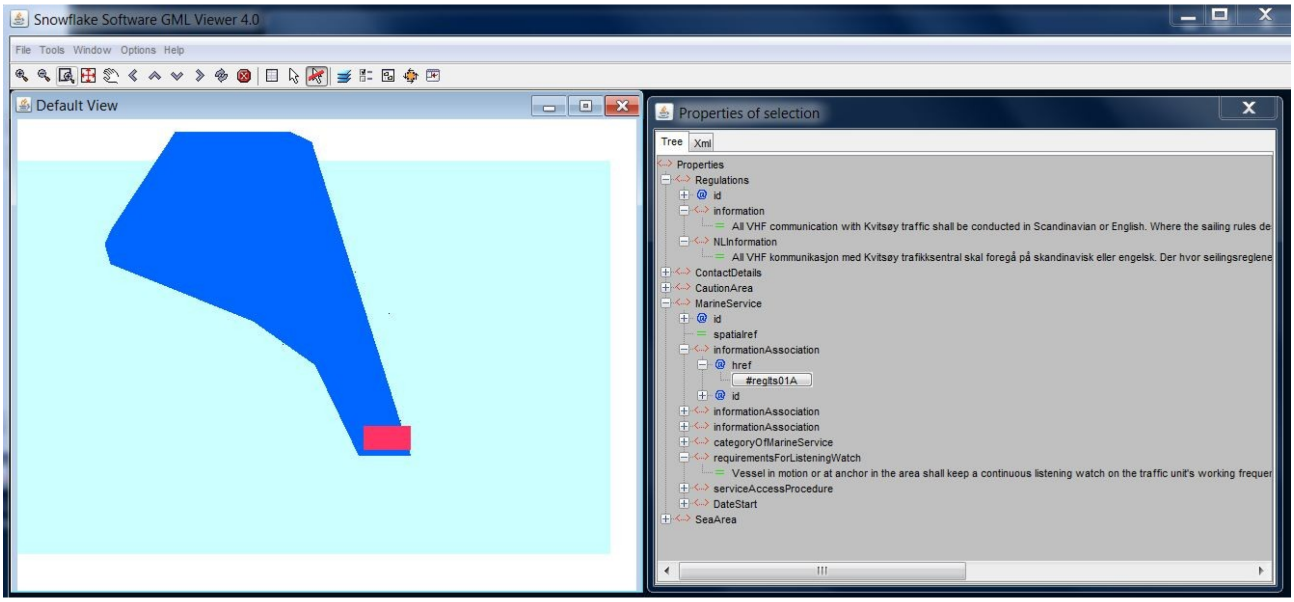


Figure 8. Area features and XML for NHS data set

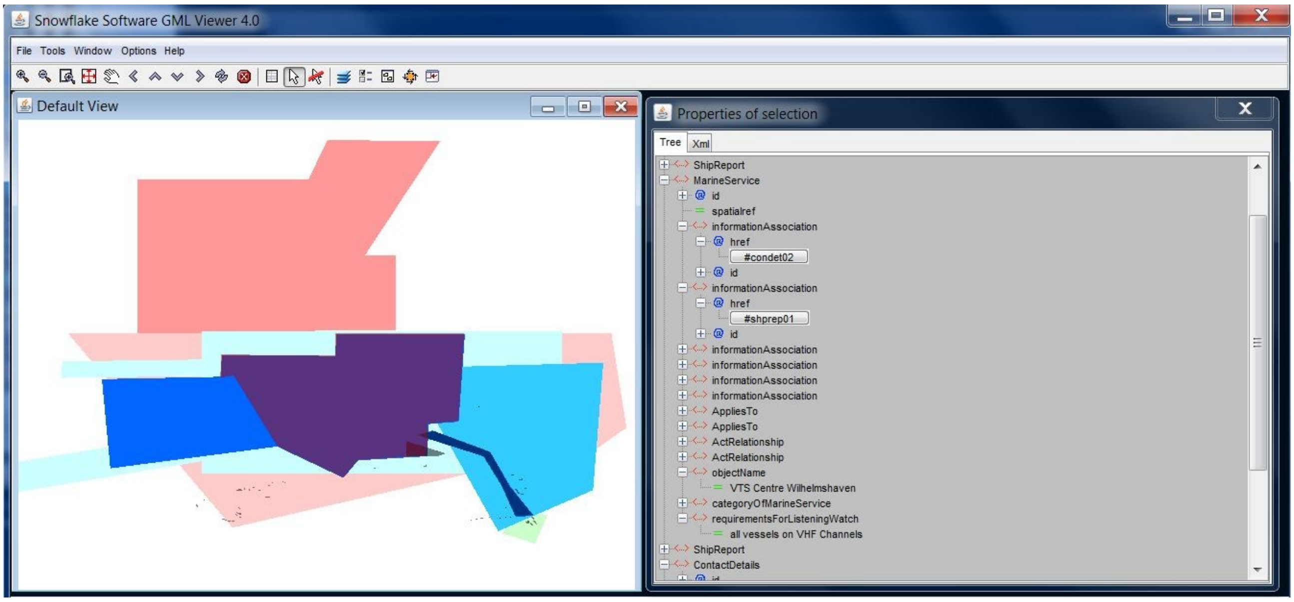


Figure 9. Selected area features and XML for BSH data set

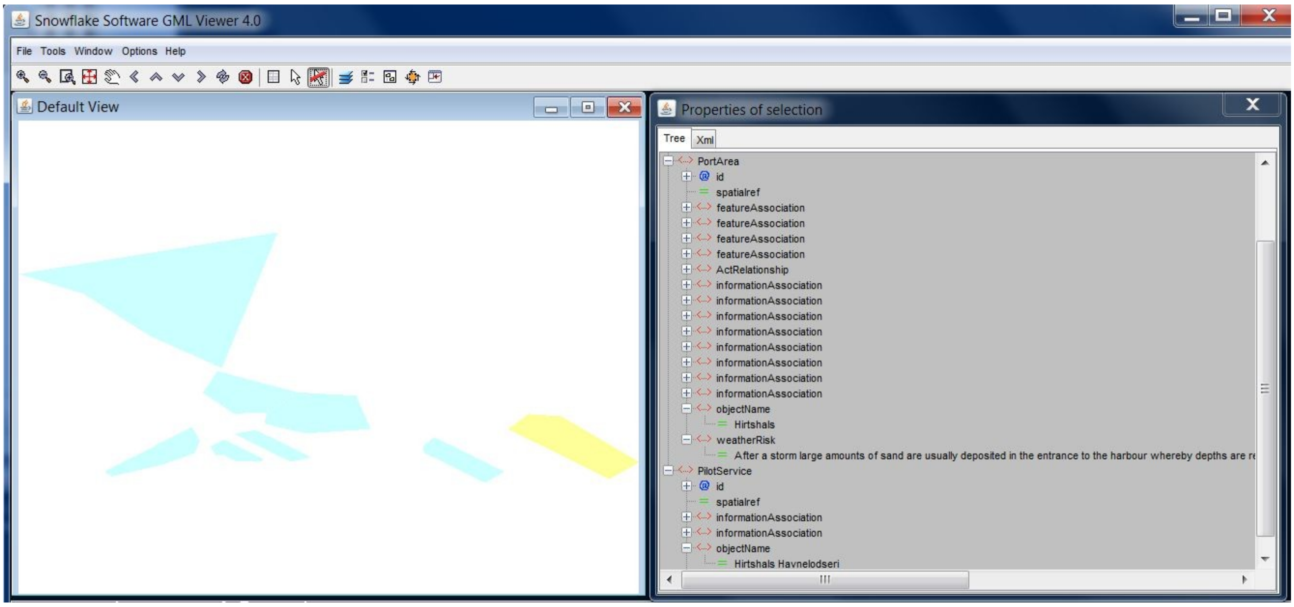


Figure 10. Area features and XML for KMS data

7. Results

The characteristics, volume and populations of the three data sets are summarised in Table 1.

Table 1. Volumes of converted datasets

Characteristic / Producer	BSH	NHS	KMS
Feature types used	32	6	5
Information types used	9	5	5
Attribute types used	96	41	19
Feature instances (geographic objects)	141	17	15
Information object instances	106	24	14
Complex attributes instances	50	13	2
Simple attribute instances	817	174	69
Feature associations	19	0	4
Parametric[2] information associations	54	11	2

[2] A parametric association is an association with one or more attributes which provide a property of the association, such as whether a class of vessels is mandatorily subject to a regulation or exempted from it. The vessel class is described by an information object at one end of the association, while the regulation is described by an information object at the other end. A parametric association corresponds to an association class in UML.

Ordinary (non-parametric) information associations	145	8	25
Size of XML dataset	163 KB	22.8 KB	20.0 KB

The populations and distributions of features and information object types for each data set are shown in Figure 11 through Figure 18.

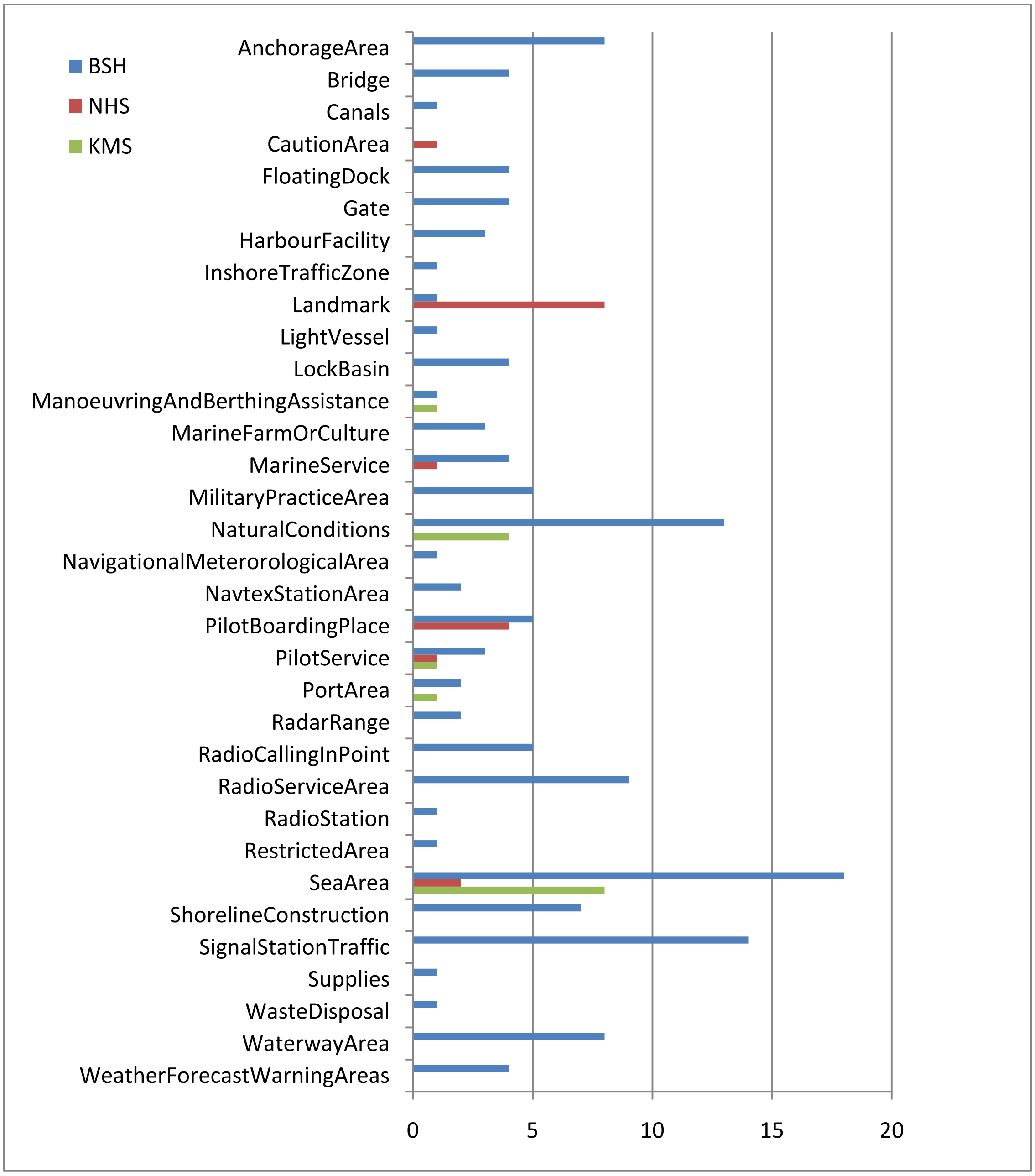


Figure 11 Comparative feature object populations

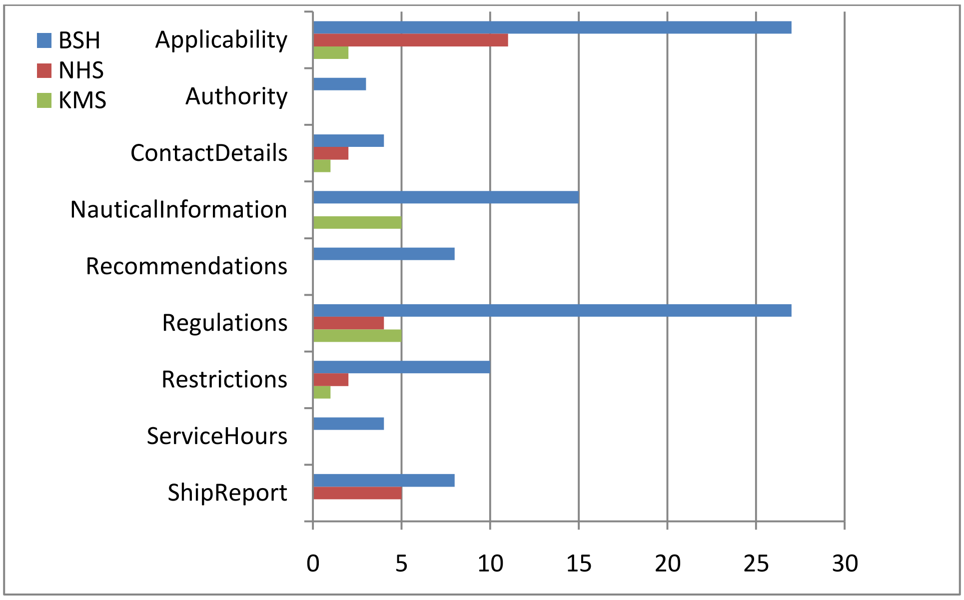


Figure 12 Comparative information object populations

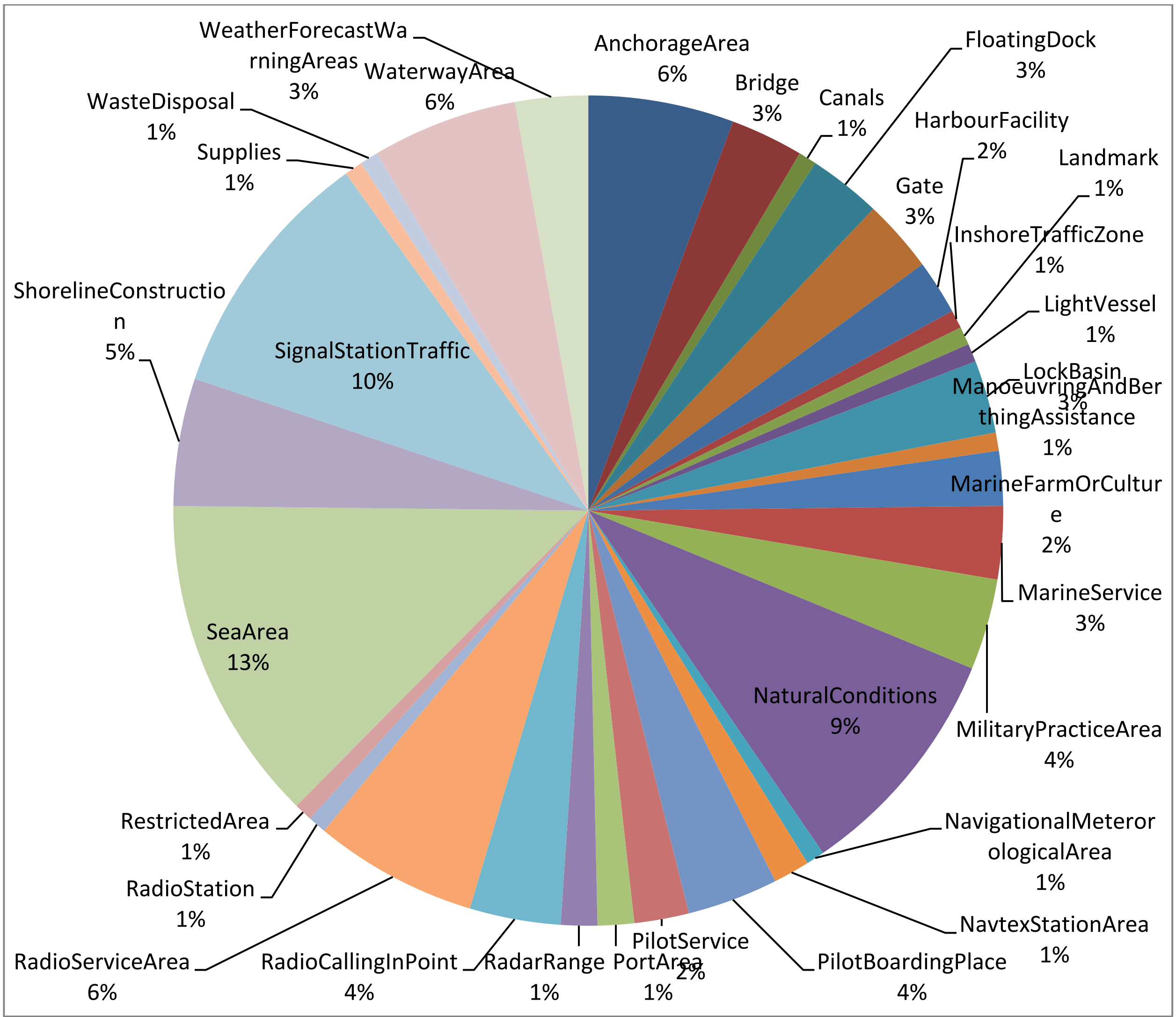


Figure 13. Distribution of feature types in BSH dataset

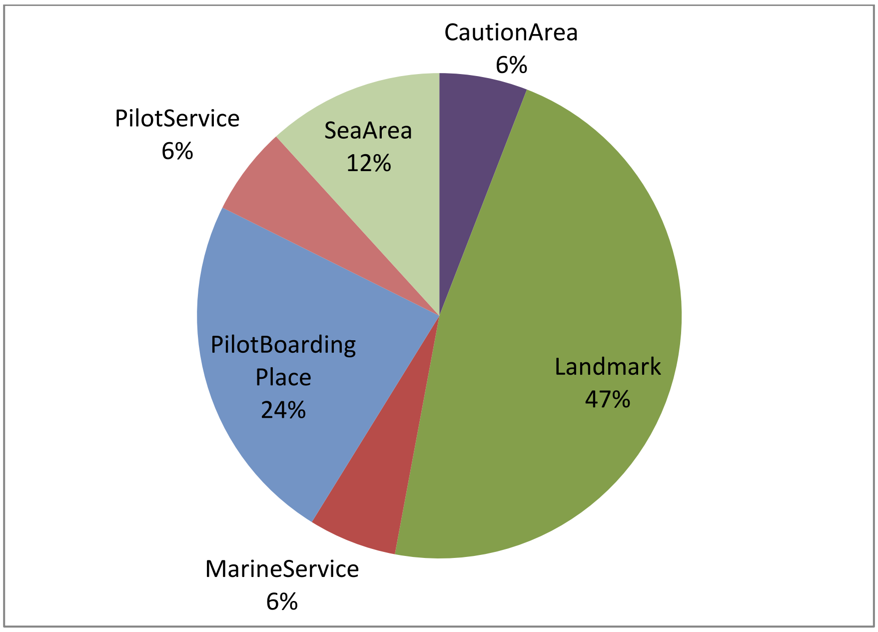


Figure 14. Distribution of feature types in NHS dataset

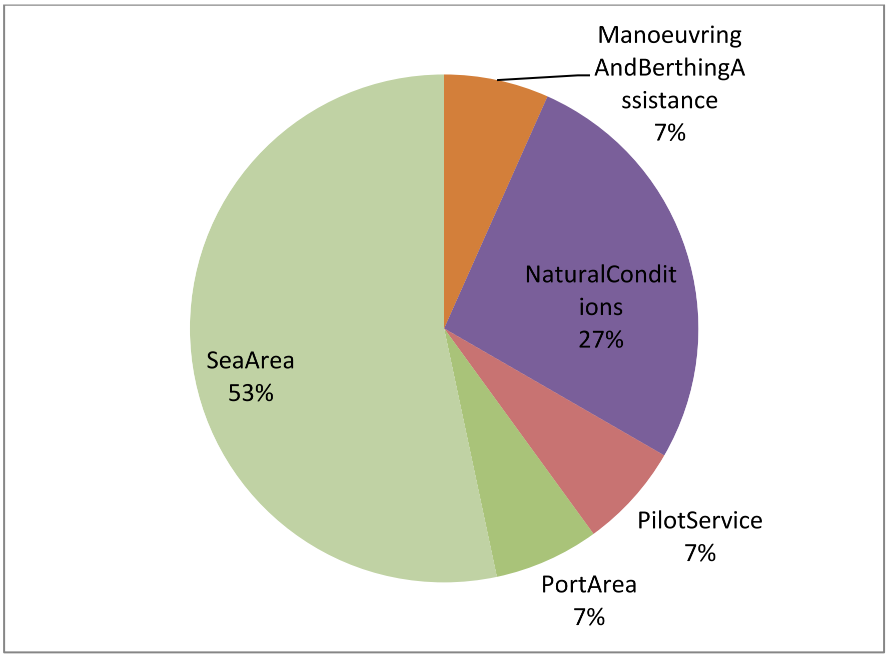


Figure 15. Distribution of feature types in KMS dataset

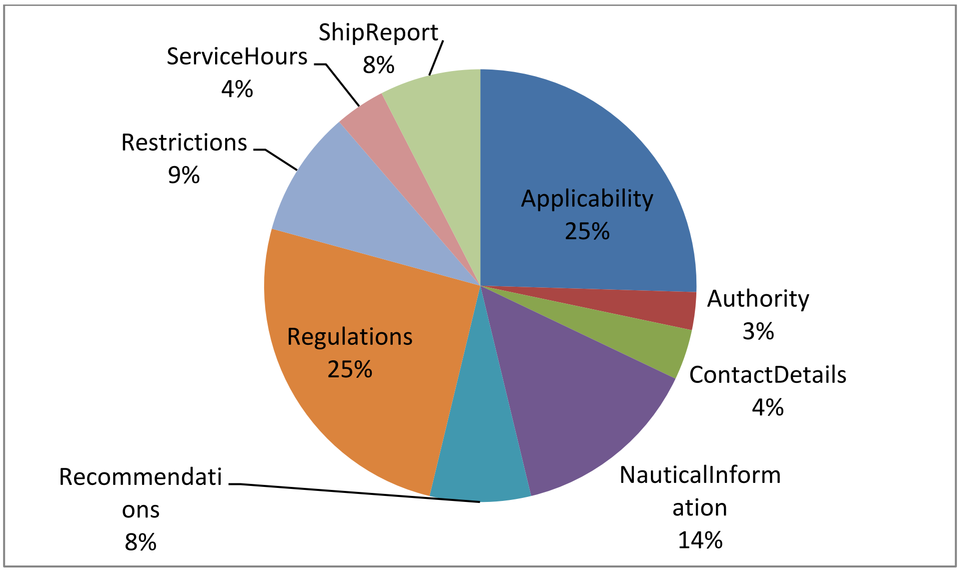


Figure 16. Distribution of Information types in BSH dataset

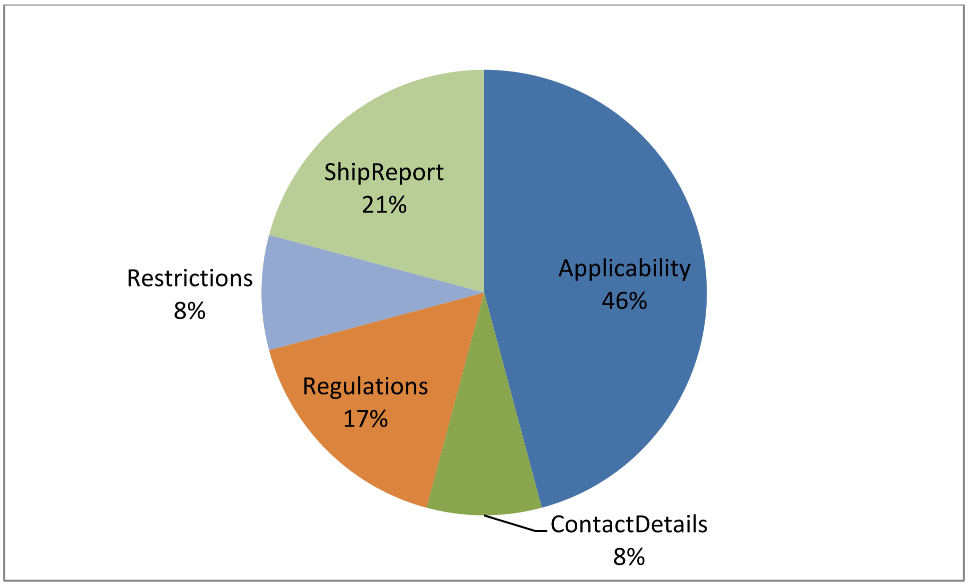


Figure 17. Distribution of information types in NHS dataset

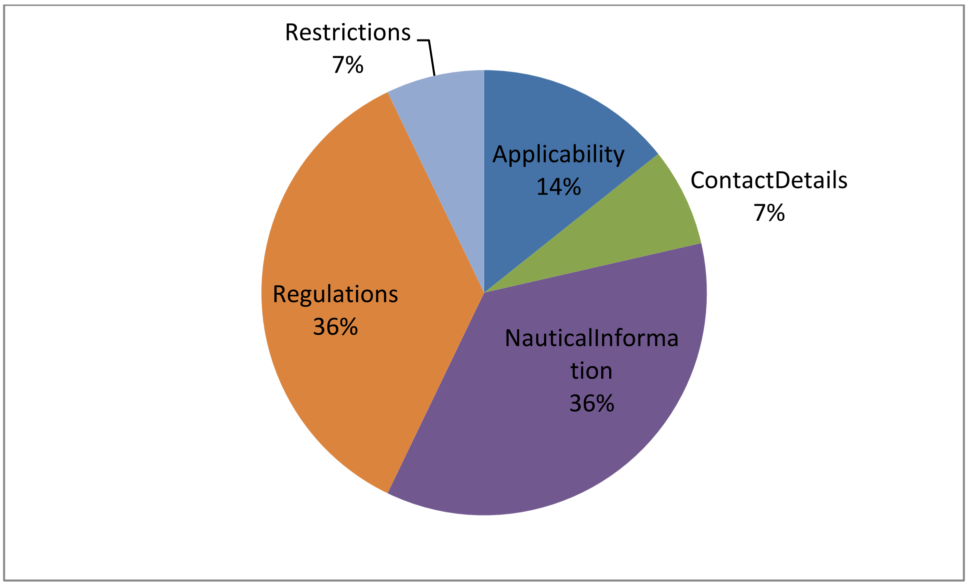


Figure 18. Distribution of information types in KMS dataset

8. Analysis

The key fact demonstrated by this exercise is that it is possible to convert a widely variant set of source documents as described in Section 3 into a common data model, indicating that it is indeed possible to harmonise nautical publications information using the SNPWG model. All converted data will be used in a common product (the digital routing guide), which will be a further test of the harmonisation concept by providing harmonised presentation of the information, as compared to harmonized data.

The volume of data varied widely. This is attributed to the difference of size of the area mapped, the complexity of marine features in the regions mapped (for example, there is no VTS service not traffic separation scheme in the near vicinity of Hirtshals). Another consideration is differences about the scope of coverage of a digital routing guide.

Data classification resulted in the discovery of several data mismatches between ENC data and nautical publications information. This suggests that the move to digital nautical publications will create a need to harmonize the publications with charted data, which will increase the workload on hydrographic offices until smarter processes using the “store once, use many times” principle are implemented.

The difficulty of converting certain kinds of nautical information (especially rules and regulations) using a data model conforming to the S-100 framework indicates that the S-100 standard needs to be extended to accommodate more kinds of information than it was originally designed for.

All three mappings made significant use of *SeaArea* objects as loci for attaching information objects. Excluding objects attributed to location-specific considerations (e.g., the port services and harbour facilities, landmarks) the BSH and NHS datasets both define about 20% of their feature objects as SeaArea objects.

All three mappings made significant use of vessel class definitions ("*Applicability*" objects) to define the class of vessel to which regulations, restrictions, or nautical information applies. All three mappings made significant use of Regulations information objects. The BSH and KMS datasets also made significant use of *NauticalInformation* information objects. This is as expected since much of the content of the source material consists of rules and regulations, and all hydrographic offices consider it important to communicate rules and regulations to mariners.

Statistical comparison of the three datasets (Table 2) shows variations in size and numbers, attributed to the difference in areas and complexity of information, as well as differing philosophies about what a digital routing guide should contain. There are, however, detectable commonalities.

1. The ratio of features with information associations to all features is relatively similar (0.51, 0.67) for the BSH and NHS datasets, suggesting that deciding whether a feature should be included in the dataset depends in part but not wholly on whether the encoder desires a geographic locus for an information object. The ratio of information objects to features with information associations is close for the same producers (1.47, 1.4).
2. Information can be included in the data set in the form of an information object or an <information> tag in a feature object. The ratio of information chunks to information objects (the "capture metric") is similar for all three producers, consistent with the use of a common data model.
3. Ratios for information chunks to information objects for "traditional" and "extended" routing guide objects are quite similar (the last two rows of Table 2) suggesting that when differences in coverage philosophies are accounted for, the end result is quite consistent across different producers and levels of detail.

Table 2. Statistical comparison of datasets

Criterion / Producer	BSH	NHS	KMS
All geographic features	141	17	15
Unique Information objects	106	24	14
Features with information associations	72	6	10
Ratio of features with information associations /all features	0.51	0.35	0.67
Ratio of information objects/features with information associations	1.47	4	1.4
Total chunks of information	127	31	18
Ratio of information chunks/features	0.90	1.82	1.20
Ratio of information chunks / information objects	1.20	1.29	1.28
Vessel class objects	27	11	2
Regulation / recommendation / restrictions / nautical information objects	60	6	11
Features which have information associations or other information	85	13	14

Ratio of chunks of information / features with information chunk	1.49	2.38	1.29
Traditional routing guide features[3]	97	17	14
Information objects associated with traditional routing guide[4]	122	24	14
Ratio of information objects/traditional routing guide features	1.26	1.41	1.00
Total chunks of information for traditional routing guide features	143	31	20
Information capture metric for traditional routing guide features	1.19	1.29	1.43
Information capture metric (all features)	1.35	1.29	1.43

9. Conclusion

This project demonstrated the conversion of unstructured textual information from nautical publications for Denmark, Germany, and Norway into structured data conforming to the IHO SNPWG data model, and produced XML datasets for the three ports. The project used the information requirements of a hypothetical digital mariner's routing guide to drive the selection of data from nautical publications. This phase of the project demonstrates that nautical publications information for disparate ports from the abovementioned North Sea nations can be harmonized by means of conversion into a single data model.

[3] Features which would appear in a traditional paper routing guide. Excludes port services, berthing, etc.
[4] Counts objects associated with multiple features as different objects

COPYRIGHT

Certain materials in this document are used by permission of copyright holders, including but not limited to the following:

Copyright BLAST, 2010

Copyright Interreg IVB North Sea Region, 2010

Copyright Danish Maritime and Safety Administration, 2010

Copyright Danish National Survey and Cadastre (KMS), 2010

NOTICE OF USE AND DISCLAIMER

Permission to use, copy, and distribute this document is hereby granted, provided that you protect and disclose all copyright holders and the entire text of this NOTICE. No right to create modifications or derivatives of this document is granted pursuant to this license.

THIS DOCUMENT IS PROVIDED "AS IS," AND COPYRIGHT HOLDERS MAKE NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED. COPYRIGHT HOLDERS WILL NOT BE LIABLE FOR ANY DIRECT, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF ANY USE OF THE DOCUMENT.

The brands, trademarks, and copyrighted material in this document may NOT be used in advertising or publicity pertaining to this document or its contents without specific, written prior permission of the copyright holder.