Hydrographic Dictionary Working Group Meeting
London (25 – 26 July 2017)

Proposals for a New Structure and Database Application

Submitted by:	IHO Secretariat
Executive Summary:	The Hydrographic Dictionary (S-32) is one of the oldest IHO publications. It is referenced by most IHO standards. This paper proposes a database model for the hydrographic dictionary content.
Related Documents:	All IHO S-10X Publications
Related Projects:	All HSSC WGs and PT's.

Introduction

1. The (S-32) Hydrographic Dictionary (HD) is one of the oldest IHO publications. It contains terms and definitions that are referenced by many IHO publications. It is also an important cross language reference. Currently the HD is available in English, French and Spanish, and includes more than 6200 items.

2. The HD is the primary source of reference for "real world features[1]" in the IHO S-100 Geospatial Information Registry. Each item in the IHO Registry is described by a unique, clear and unambiguous **definition**. Each definition must have an associated term (or terms) and a unique identifier[2]. The unique identifier should be used as the mechanism for cross referencing items in the HD, Registry and other IHO publications. There can be only one unique identifier per definition. The unique identifier should also be the mechanism used for linking equivalent items across language domains. (See examples at Annex A).

3. This paper proposes that the HD content is used to populate a dictionary database with certain Registry characteristics. These include: the use of unique identifies for records; and record persistence, i.e. items are never deleted from the database. Registered items have a status of either "Proposed", "Valid," "Superseded" or "Retired". References to a registered HD item (i.e. included in IHO publications) will always locate an item, irrespective of whether it is "Valid", "Superseded" or "Retired". (Superseded and Retired items will have a pointer (link) to the associated "Valid" Item). The Registry governance principals (see S-99) ensure that items must go through a formal review and acceptance process before being entered into a register.

What's the difference between Concepts, Terms and Definitions?

4. ISO 1087-1:2000 defines a "Concept" as "a unit of knowledge created by a unique combination of characteristics". A "Term" is defined as "a verbal designation of a general concept in a specific subject field". A "Definition" is described as "a representation of a concept by a descriptive statement which serves to differentiate it from related concepts". A "Definition" does not describe a "Term" or other designation, and it should not contain characteristics that belong logically to superordinate or subordinate concepts.

[1] Non abstract features such as buoys, beacons, rocks etc…
[2] In the common language, one definition can correspond to more than one term (synonyms). Any term may correspond to more than one definition (polysemy).

Proposed Data Model for the Hydrographic Dictionary Database.

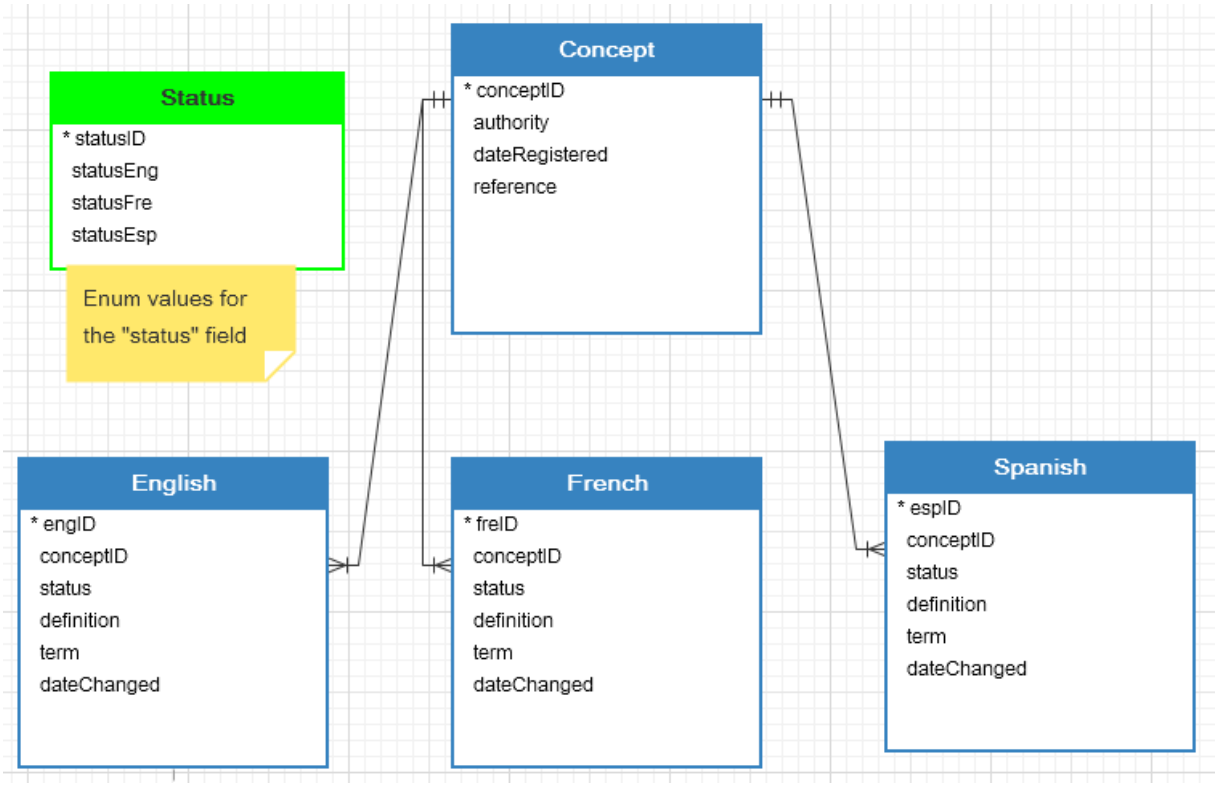


Figure 1 – Proposed HD Database Model

5. See test database application at: http://www.iho-ohi.net/S32/

Note: The current S-32 database model includes "English", "French," "Spanish" and "Concept" tables. The "Status" table (shown in green) is used to populate a list of enumerated values, i.e. a dropdown list of fixed values. The relationship and cardinality between the tables is shown in "crows foot" notation which is illustrated in Figure 2 below. A description of the tables/fields is provided in Tables 1 to 3.

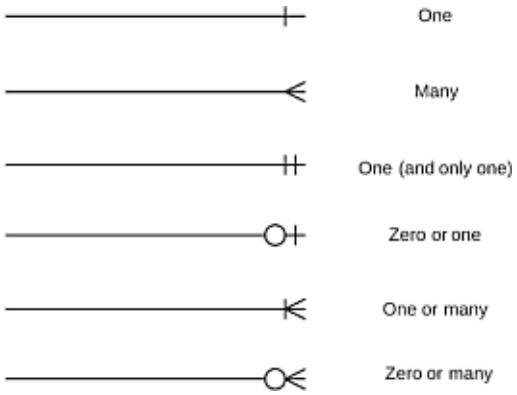


Figure 2 – Crows Foot Notation

6. Database table: Concept (parent table that includes common metadata)

Name	Description	Note
conceptID (PK)	Unique identifier	Mandatory field (Is linked to by English:conceptID, French:conceptID and Spanish:conceptID
Authority	Originator of the concept	date field
Reference	Reference to the source information – if appropriate	text

Table 1

7. Database:table: English (Same table structure is used for all language database tables).

Name	Description	Note
engID (PK)	Unique identifier for the table	
conceptID (PK)	Unique identifier links to Concept:conceptID	Mandatory
Status	Status of the concept	Enum list from Status table (1 = Proposed, 2 = Valid, 3 = Superseded, 4 = Retired)
Definition	Concept definition	Text
Term	term	Text
dateChanged	Date when the producer was last changed	Text

Table 2

8. Database Table:Status (To be used as enumerated values)

Name	Description	Note
statusID (PK)	Unique identifier for the table	Mandatory
statusDescriptEng	Description in English	Eng: Proposed, Valid, Superseded …
statusDescriptFre	Description in French	Fre: Proposed, ….
statusDescriptSpa	Description in Spanish	Spa: ….

Table 3

Current Status of HS Content

9. The current (WIKI) version of the HD has not been regularly maintained. In 2015 the IHO Secretariat Project Officer from Peru undertook a project to revise the content of the three language paper document editions, and also introduce a numeric identifier to link equivalent terms across the three languages. The output from this endeavour was an Excel document containing a complete match between the English and Spanish terms and definitions, but only about 80% of the content for the French items. The items in the spreadsheet include the following;

- Unique identifier (numeric)
- Term (text)
- Term (repeated) + Definition.

10. In order to produce the test database application, the IHO Secretariat also included the terms and definitions from the Glossary of ECDIS-Related Terms (formerly S-52, Appendix 3). A script was developed to delete the repeated term (text) in the Definition column, however there is a need to review the Definitions in order to clean up anomalies. Some additional metadata attributes were also added.

Recommendations

11. The WG is invited to consider the issues raised in this paper and discuss;

- whether to continue with a WIKI structure or adopt a Registry (type) structure;
- how to complete the missing content (especially the French content);
- how to clean up anomalies in the data (especially the text in the “Definition” column);
- whether the proposed database model is suitable;
- what additional tables / fields should be added;
- what additional tests need to be carried out.

12. Furthermore the WG are invited to consider how any new HD application will be structured and managed in order to support the IHO Registry, taking into account the issues raised in paragraph 2.

Justification and Impacts

13. The current WIKI Hydrographic Dictionary application is not suitable for managing multilingual dictionary terms and definitions. It does not make provision for item persistence, and does not enable linking of equivalent

items across language domains. Any future HD application should cater for these, and should facilitate management / governance procedures similar to those described in S-99.

Action Required

14. The WG is invited to consider the proposal herein and agree a way forward for developing a new online Hydrographic Dictionary application that will support the IHO Registries Concept Register and dictionary reference in other IHO publications.

Annex A

Example HD Database Application

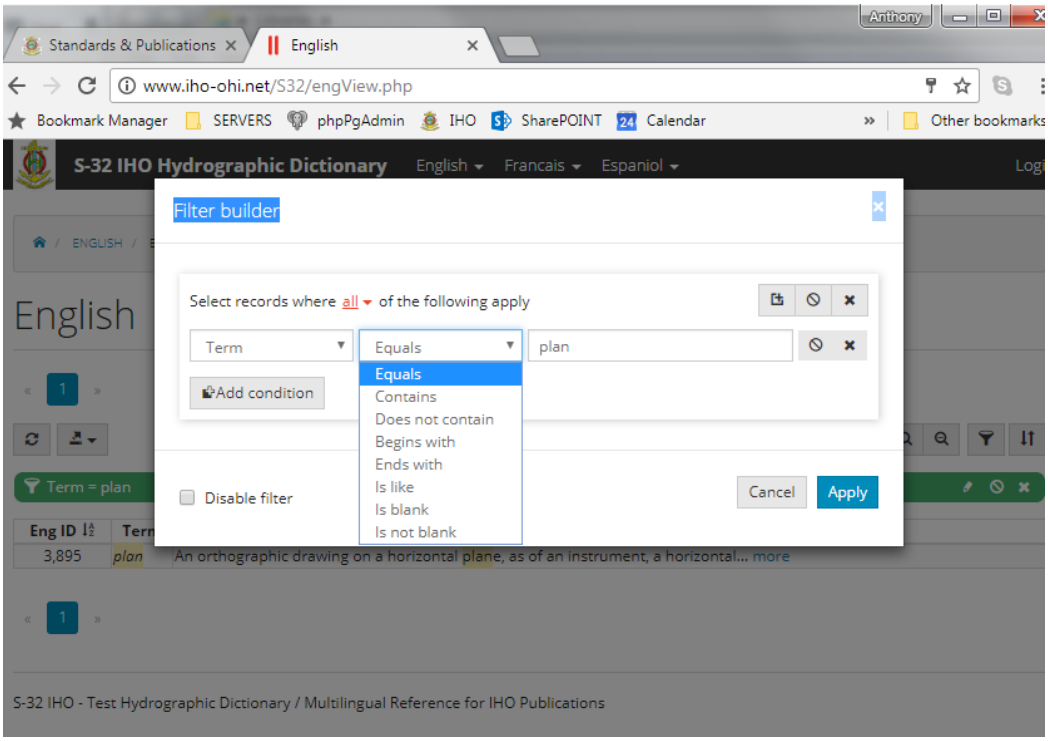
An example database web application has been setup in order to demonstrate some possible functions.

Views

Database "Views" show English, French and Spanish content; including language combinations e.g. English – French, French – Spanish, Spanish – English etc… (Please see http:www.iho-ohi.net/S32/). Other languages can easily be added provided that they include matching unique IDs, terms and definitions).

Filters

Basic and advanced filter (search) mechanisms have been set up. The basic "Quick search" will identify all matching items in both the "Terms" and "Definitions". The advanced "Filter builder" enables searches to be made on different types of matching criteria, e.g. Equals, Contains, Does not contain, Begins with, Ends with etc…. These have only been programmed for the English, French and Spanish views only.



Filter builder

Referencing Database Items

The HD is an important resource for many IHO publications and the Registry. It should be possible to reference individual HD items in the database. This could be done using the database application URL as shown below. Note: there may be more sustainable ways of doing this.

Example reference in S-4 Ed 4.7.0

B-254 REFERENCES TO OTHER CHARTS

Hydrographic offices should include on their …….

Note: Insets, including continuation insets and large-scale plans ….. larger chart. A plan is a large scale inset of a nautical chart (for example a port plan). For more detailed definitions, see **the Hydrographic Dictionary, S-32**.

The unique ID for the “plan” item (concept) is 3895.

(This reference could be; “see ***the Hydrographic Dictionary, S-32 – ID 3895”)*** or (“*see* ***the Hydrographic Dictionary, S-32”***)

The URL’s pointing to database items 3895 are;

English - http://www.iho-ohi.net/S32/english.php?operation=view&pk0=3895

French - http://www.iho-ohi.net/S32/French.php?operation=view&pk0=3895

Spanish - http://www.iho-ohi.net/S32/Spanish.php?operation=view&pk0=3895

Similar references (links) to HD items could be included in the Registry.