3rd S-100WG MEETING

Singapore – 10-14 April 2018

Paper for Consideration by the S-100 Working Group (S-100WG)

Proposed Structure of the IHO Geospatial Information (GI) Registry

Submitted by:	IHO Secretariat.
Executive Summary:	Continued practical use of the IHO Geospatial Information (GI) Registry since it became operational in October 2016, and decisions at S-100WG2 to add an overarching Concept Register to the structure have prompted further discussion within the IHO Secretariat.
Related Documents:	1. S-100WG2-9.2 – *Summary of Activities of the IHO GI Registry (rev1)*. 2. S-100WG2-9.3A – *Proposed Extension to the IHO Registry*. 3. S-100WG2-9.5 – *Review of IHO GI Registry Content (rev1)*. 4. *S-100WG2 Minutes and Actions*. 5. S-100TSM5-04.3 – Proposed Structure of the IHO GI Registry.
Related Projects:	Development of the IHO Geospatial Information Registry.

Introduction / Background

1. The IHO Geospatial Information (GI) Registry became operational in October 2016. From this time a significant amount of work has been done in registering new items in the Feature Concept Dictionary (FCD) Register, as well as performing overall Registry management functions and discussing possible improvements that can be made in the Registry structure based on the experience gained so far.

2. At the S-100WG2 meeting (Genoa, Italy, 15-18 March 2017) a number of papers were submitted to the meeting related to the Registry structure and activities, which resulted in several decisions and actions being taken, including a revision of the Registry structure to include a Concept Register; a review of the content of the Registry; and perform a resultant review of IHO Publication S-99 – *Operational Procedures for the Organization and Management of the S-100 Geospatial Information Registry*.

3. In the course of progressing these actions from S-100WG2 and continuing the day-to-day activities of the Registry, there have been further discussions within the IHO Secretariat as to how the Registry structure may be further refined. In order to best summarize the outcomes of these discussions, it was decided to develop a schematic diagram and some supporting explanatory notes to capture these outcomes.

4. This paper is a revision of paper S-100TSM5-04.3 – Proposed Structure of the IHO GI Registry (September 2017).

Analysis / Discussion

5. The schematic diagram reflecting the thoughts of the IHO Secretariat and showing the overall refined structure of the IHO GI Registry, incorporating an indication as to the flow of activities from the submission of proposals from Submitting Organizations through to the implementation of these proposals in a published Product Specification, is included as Annex A to this Paper. Explanatory and supporting notes for the diagram are also included. These notes have been supplemented by comments from members of the IHO GI Registry Project Team. It is important to note that, in the development of this proposed revised structure, research has been conducted with the implementation of other Geospatial Registries, principally the EU INSPIRE Registry (http://inspire.ec.europa.eu/registry).

6. In addition to the inclusion of the fundamental Concept Register as agreed at S-100WG2, the following aspects of the refined Registry structure should be noted:

- Within the Feature Data Dictionary Register, each S-100 based Product Specification will normally have its own Domain, which effectively acts as the "sandbox" for Product Specification development.
- In order to ensure consistency of the composition of enumerated and "open enumeration" Codelist type attributes, it is proposed that an "Enumerate Register" (and possibly a "Codelist Register") be implemented.
- The Feature Catalogue Builder and Portrayal Catalogue Builder are the tools used for extracting concepts from the Concept and Enumerate/Codelist Registers for the application of binding of concepts, conformant with the relevant Application Schema (data model), within the Domains; and for creation of draft and finalized Feature and Portrayal Catalogues for inclusion in Product Specifications.
- The proposed relationship between the Registry and the Hydrographic Dictionary (database) was discussed at the HDWG1 meeting (London, UK, July 2017). It is acknowledged that such a relationship is required, and discussion as to how this could be implemented is ongoing.

7. Also note that the main focus of the schematic at this stage is the relationship between the Concept Register and the Feature Data Dictionary Register. Further work is required to develop the structure and relationships for the Portrayal Register. It anticipated that this will occur as further experience is gained in the current Registry structure.

8. In addition to the presentation of this proposed structure at the S-100TSM5 meeting in September 2017, the schematic and supporting notes have been circulated to the IHO GI Registry Project Team members for their initial thoughts. It is expected that further refinement of the Registry structure will be required as more knowledge/experience in some of the concepts that have been suggested for implementation in the Registry is gained, and further discussed. This includes, but is not limited to:

- Investigation of the requirement of the structure and content of the Concept Register being flexible enough to allow Product Specifications to make the necessary distinctions (essentially equating the Concept Register to a thesaurus). Investigations may be carried out in the use of the architecture of existing lexical databases, for example Wordnet; or ontology (in RDF (https://www.w3.org/standards/techs/rdf), OWL (https://www.w3.org/TR/owl2-overview/), or SKOS (https://www.w3.org/TR/skos-primer/), specifications for which have been published by ISO or W3C). Alternatively, the Concept Register may retain a relatively "flat" structure, with the required flexibility introduced through explicitly allowing, in the architecture and guidance documents, for product specifications and the data dictionaries to make different types of derivations from the Concept Register (especially refinements, and specializations or partitions);

- Further to the above, investigation is required as to the implementation of “scopes” of concepts, essentially through the use of Namespaces, in order to define a context (as required) for registered items. Is it required to implement this at the Concept Register level, or is it sufficient to implement at the Data Dictionary Register level (and possibly within the Enumerate/Codelist Register(s))?;
- The notion of limiting the scope of the Domains to a single Product Specification requires further discussion. This is likely to be troublesome for implementers. It may also be troublesome for Product Specifications that share parts of their application schemas. There is a need to develop a middle ground between total independence and total integration of different product specifications;
- Discuss the merits of the introduction of the concept of a « datatype » in order to resolve issues of duplication of items at the Concept Register level. This is also related to the proposed introduction of Enumerate/Codelist Register(s), since a set of enumerates (listed values), is effectively a « datatype »;
- Discussions related to the issues raised in associated papers submitted to S-100WG3 related to the contents of the current Feature Concept Dictionary Register; and the development of a set of conventions and guidelines for IHO GI Registry content.

9. It is anticipated that the changes to the Registry structure and the resultant impact on the contents of the (current FCD) Concept Register have the potential to significantly impact all S-100 based Product Specifications currently in development. In order to provide as much transparency as possible and allow stakeholders to provide input, it is recommended that an “IHO GI Registry Workshop” be conducted. Attendance at such a Workshop would be reserved for the IHO GI Registry Project Team members; representatives of groups developing S-100 based Product Specifications (Submitting Organization and Domain Control Body representatives, WG/PT Chairs); and relevant subject matter experts.

Conclusions

10. The IHO GI Registry structure as shown in this Paper reflects the outcomes of discussions within the IHO Secretariat based on experience and observations from the day-to-day administration and management of the Registry. This structure has been refined incrementally as discussions have progressed in order to resolve issues that have risen from Registry use and the ongoing review of the content of the FCD Register.

Recommendations

11. S-100WG to consider the proposed structure and supporting notes at Annex A; and the related additional discussions required as summarized in paragraph 8, and provide comment and further input, noting that this structure reflects principally (at this stage) only the thoughts of the IHO Secretariat based on its experience with the administration and management of the Registry, supported by some initial observations from the IHO GI Registry Project Team. In order to facilitate a focussed discussion on the issues with the Registry structure and its content, it is recommended that a dedicated IHO GI Registry Workshop be convened.

Justification and Impacts

12. Refinement of the structure of the IHO GI Registry is required in order to optimize the management and administration processes of the Registry; consolidate, and improve management of, Register content; and provide an appropriate environment for S-100 based Product Specification developers to do their work without unnecessary Registry oversight.

13. Impact will be principally on the Registry development team at KHOA and IHO Secretariat staff. However, dependant on the outcomes of further discussion, there may also be significant impact on S-100 based Product Specifications in development.

Action required of S-100WG

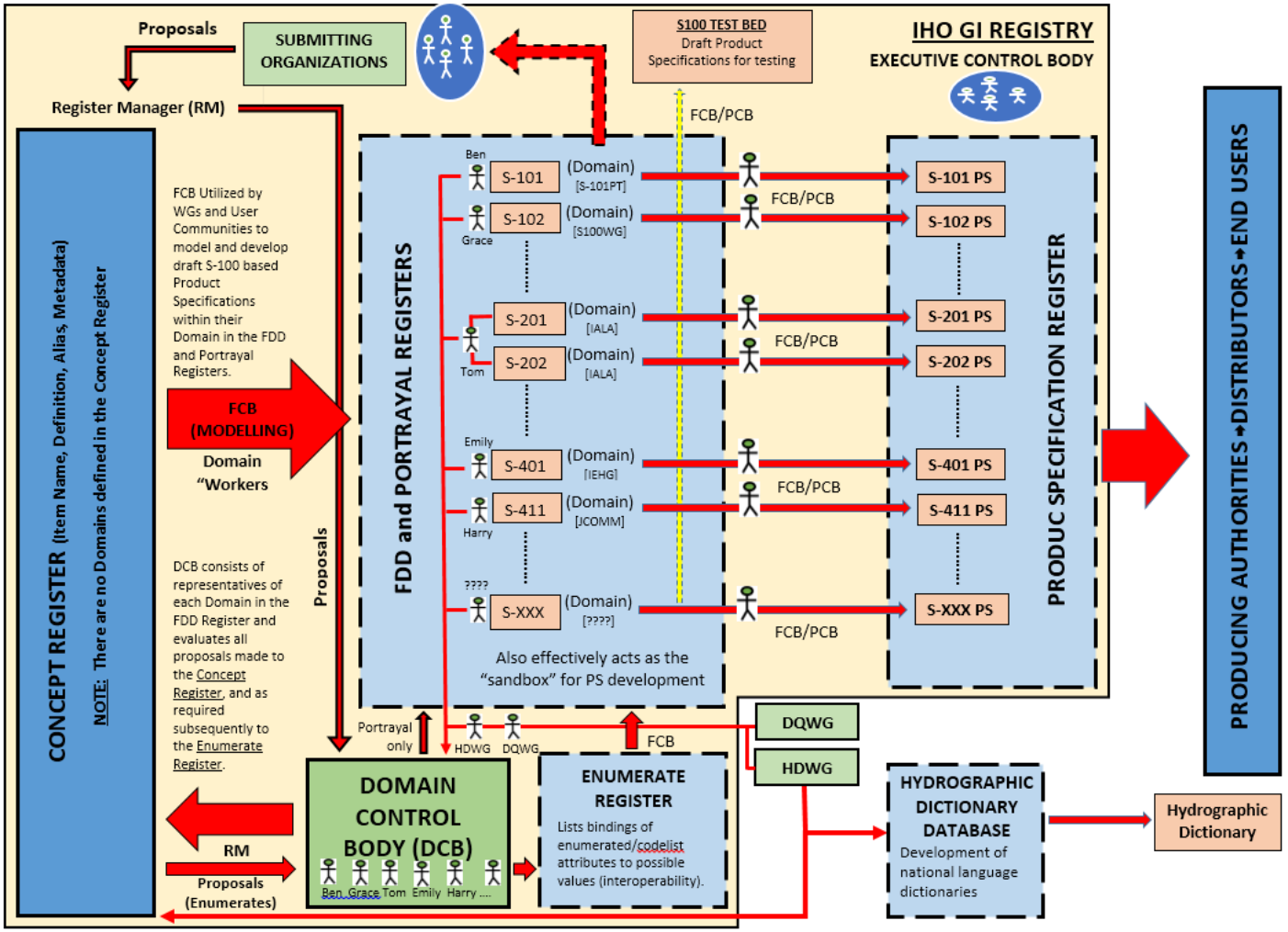
14. The S-100WG is invited to:

    a. **Note** this paper.

    b. **Provide** ongoing comment and feedback to the IHO Secretariat as refinement of the Registry continues.

    c. **Consider** the possible further discussion required for refinement of the IHO GI Registry structure, as summarized in paragraph 8, and endorse the continuation of this discussion, through the hosting of an IHO GI Registry Workshop.

ANNEX A to Paper S-100WG3-06.2.2



IHO GEOSPATIAL INFORMATION (GI) REGISTRY STRUCTURE

Explanatory and Supporting Notes

Concept Register:

- The Concept Register is effectively the "source" register from which all hydrographically-relevant concepts are drawn for modelling in S-100 based Product Specifications.
- A single instance only of each concept exists in the Register. Each concept must be unique (that is, no two concepts can be interpreted to describe the same "real world" entity), and is described by Item Name; Definition (with supporting metadata); Unique ID; Alias(s) (if any); Status (Valid, Invalid (or Not Valid), Superseded, Retired); Lineage and Maintenance Metadata; and a flag to identify whether the concept is included in the IHO Hydrographic Dictionary.
- The Concept Register is not partitioned into separate Domains. In other words it is "domainless". However, assessment of proposals to the Concept Register is done by the Concept Register Domain Control Body (DCB), which consists of representatives of each of the Domains contained in the Data Dictionary Register (and HDWG, DQWG) – see notes on DCB below.
- IHO GI Registry Process:
    - Submitting Organization submits a proposal to the Register via the IHO GI Registry interface;
    - Proposal is assessed by the Register Manager for completeness and possible duplication with items already registered. If suitable, the proposal is forwarded to the Concept Register Domain Control Body. [If considered to be not suitable, the proposal is "rejected" and returned to the Submitting Organization for further rework/resubmission or withdrawal based on Register Manager comments.];
    - Proposal is assessed by the Concept Register Domain Control Body for suitability and possible impact on Product Specification(s) under the individual members' area of expertise. If approved, the proposal is forwarded to the Register Manager for incorporation in the Register. [If rejected, the proposal is forwarded by the Register Manager back to the Submitting Organization for rework; or appeal by the Submitting Organization to the Executive Control Body.];
    - Register Manager commits the approved change to the Concept Register, and the Submitting Organization is notified of the change, from which time the change is available for use in the Feature Data Dictionary Register and/or the Enumerate Register.

Feature Data Dictionary (FDD) Register:

- The Data Dictionary Register is partitioned into Domains. Each Domain will generally correspond to a single S-100 based Product Specification. It has been proven that having multiple Product Specifications being derived from a single Domain causes problems within the Domain as the possibility exists that multiple instances of a single concept modelled in different ways may be required in the Domain dependant on the requirement of each Product Specification. Multiple Product Specifications within a single Domain should be considered only where these Product Specifications share all (or most) of a single Application Schema; or the Application Schema for a Product Specification is essentially a "subset" of the Application Schema for another Product Specification.
- Concepts are drawn from the Concept Register by nominated representative(s) from the relevant IHO Working Group or User Community, utilizing the Feature Catalogue Builder (FCB), into a Domain within the Register. Within the Domain, Feature Catalogue development (assign geometry; type; binding; multiplicity) based on the Application Schema for the Product Specification is done.
- There is no overarching IHO GI Registry structure or process governing how the development work within a Domain is managed. This is the responsibility of the Working Group or User Community that is developing the Product Specification. There is no requirement for the Register Manager, Registry Manager, DCB, or ECB to be involved in the actual development of the Product Specification, except for the initial establishment of the Domain; processing new proposals from the Domain Submitting Organization representative to the Concept Register; and providing advice and guidance as required. All responsibility for ensuring a complete

Commented [TS1]: RM : I am arriving at the realization that the concept register idea should be more flexible than this document and the diagram suggest. Either the CR should allow for "**senses**" or "**scopes**" like a **lexical database or thesaurus**, or the architecture and guidance documents should explicitly allow for product specifications and the data dictionaries to make different types of derivations from the concept register (especially refinements, and specializations or partitions). Use the concept register to link the derivations, and require derivations to describe the relationship to the entry in the concept register, e.g., which item (and which sense) it relates to and the nature of the derivation (refinement, specialization).
JW: Not sure what this means – need some worked examples.
RM: Will try to put together a few slides for the S100WG meeting.

Commented [TS2]: RM : The tendency will be to reject anything that might have an impact, which would block development of new products or introduction of new maritime information domains. Adding **senses or scopes** would mitigate, though not avoid, this potential problem.
JW: Need to know more about "senses or scopes" – how is this implemented?
RM: The simplest implementation might be just to add a « senseNumber » to the item record for a term, so we can have different senses for the same term.

Commented [TS3]: RM : This is likely to be very troublesome for implementers. It will also be troublesome for product specifications that share parts of their application schemas.
JW: Have amended the wording, however need to know more about "scopes" sand "namespaces".
RM: In its simplest form, a scope is a « container » for terms (or other scopes); a namespace is more or less the same idea as « scope » but carries the connotation of uniquely identifying an item by prefixing its local identifier (in its immediate scope) with one or more names identifying the scope hierarchy in which it is defined. For example, URNs.

Commented [TS4]: RM : Introduce **scopes or namespaces**? Need to develop a middle ground between total independence and total integration of different product specifications.

Commented [TS5]: RM : The application schema (UML model) is developed first, then the feature catalogue. The project specification team cannot know which concepts are needed until the application schema is completed. It develops the application schema by a process of iterative refinement, referring to the concept register as a source (though not the sole source) of concepts within the scope of the data product. The feature catalogue is developed later. When the FDD is introduced, populating it will be an intermediate step between developing the application schema and feature catalogue.
JW: Amendments made in accordance with the above comment.

and robust process in order to produce a fit-for-purpose Product Specification are the responsibility of the governing IHO Working Group or User Community (noting however the existing approval process for IHO S-100 based Product Specifications).

- The process and participants for development and maintenance of the Product Specification can be organized by the Working Group or User Community responsible as required so as to best achieve the required end result. For example, the IHO S-101 Project Specification is being developed by a dedicated Project Team operating under the S100WG, while S-102 was developed by a very small group of subject matter experts (essentially a "one man band"), and simply reported its progress to the S100WG as required. Similarly, cooperation between Domains may be "sub-managed" by smaller cross-Domain groups in order to harmonize and optimize Product Specification development – for example the IHO Hydro "Cross-Domain Group" between the S-101 Project Team and the NIPWG. Again, it is important to note that this is not a part of the overarching administration or management of the IHO GI Registry.
- At any stage during Product Specification development, a draft product Feature and Portrayal Catalogue may be created (utilizing the Feature Catalogue Builder and Portrayal Catalogue Builder) from the Domain within the FDD Register for testing in the S-100 Test Bed. This effectively means that the Domain space within the Feature Data Dictionary Register acts as the "sandbox" for the iterative development and refinement of the Application Schema and Feature/Portrayal Catalogues for the Product Specification.
- When all requirements for the development, testing and approval of the Product Specification have been satisfied, the final Feature and Portrayal Catalogues are produced, utilizing the Feature Catalogue Builder and Portrayal Catalogue Builder, and included in the published Edition of the Product Specification.
- The published Product Specification is included in the Product Specification Register, which holds all published versions of the Specification. From this point, further development can be done in the FDD Register for the next draft of the Product Specification, as required.
- IHO GI Registry Process:
    - The Working Group/User Community applies to the Registry Manager to have a Domain assigned to them for an S-100 based Product Specification.
    - When approved by the Registry Manager, the Domain is established. The Working Group/User Community then assigns representative(s) of their group to act as Submitting Organization, Domain Control Body and Domain "Worker". The Domain "Worker" essentially has write access to the Domain for the application of the data modelling for the Product Specification, and is given access to the Feature Catalogue Builder for interface with the Concept Register so as to create draft Feature Catalogues for testing and final publication.
    - Based on draft modelling included in the Application Schema, the Working Group/User Community for which the Domain has been created extracts concepts from the Concept Register, and models the concept according to their requirements (assigns geometry, type, binding, cardinality, encoding guidance). This is done by the Domain "Worker" utilizing the Feature Catalogue Builder. The Registry interface provides a query mechanism whereby users of the Registry can enquire as to how a concept from the Concept Register has been modelled in all instances of its use in the FDD Register and within the Enumerate Register – this will assist in Product Specification development and contribute to interoperability;
    - Proposals for new or revised concepts required to the Concept Register are proposed by the Submitting Organization representative for the Domain as required;
    - As required, a draft Feature Catalogue can be extracted from the Domain, utilizing the Feature Catalogue Builder, for testing in the S-100 Test Bed.

Portrayal Register:

- It is anticipated that the Portrayal Register will be structured and function essentially as it exists in the current version of the Registry.

Enumerate/[Codelist] Register(s):

- The intention of the Enumerate Register is to provide the mechanism for ensuring consistency and interoperability between data created conformant to S-100 based Product Specifications. The Register is a

Commented [TS6]: RM : The enumerate register should define literals (enumerates, listed values) in different **scopes or namespaces**, generally corresponding to attributes. Provision should be made for hierarchies (supersets/subsets).
Enumerations (and codelists) are actually different datatypes for the purposes of implementations, data formats, and modeling and their treatment in the registry should facilitate that.
JW: Refer to highlighted NOTE in text. If the enumerate values are bound in the Register to an enumerated or codelist attribute data type, does this constitute an implementation of namespaces (for example, categoryOfSignalStationTraffic::berthing; actionOrActivity::berthing)?
RM: Yes.

"Hierarchical Register", and contains all instances where a concept from the Concept Register has been modelled in an S-100 based Product Specification as an enumerated attribute or an "open enumeration" Codelist type; and the full list of allowable enumerate codes and their values (which may also be taken from the Concept Register) for the attribute. The rationale behind the establishment of an Enumerate Register is that, if such a Register does not exist, Product Specification developers could create their own enumerate lists for the same enumerate type attribute, having different values assigned to enumerate codes. This would cause considerable problems with interoperability. NOTE: Discussion as to whether there should be separate "Hierarchical Registers" for enumerates and codelists (refer to INSPIRE Registry model), rather than a single Register, is required. An alternative is to have 2 Registers – the first being a "fixed list" Enumerate Register for enumerated lists that are stable (that is, are not intended/forecast to change); the second being an extensible "Codelist Register" containing lists that equate to an "open enumeration" Codelist that can be implemented in a Feature Catalogue as Enumeratred or Codelist type.

- As for the Concept Register, the Enumerate Register is "domainless". There must only be a single instance of any concept from the Concept Register defined as a "parent" attribute in the Enumerate Register, with all possible values (codes) as used in any S-100 based Product Specification listed against that attribute. User communities may then define a "subset" of the listed values dependant on the requirement of their Product Specification.
- All enumerated attributes and their enumerate values are derived from the Concept Register. [This may not be the case – see below for discussion required as to unique "coded" lists.]
- Data modellers working within their Domain within the Feature Data Dictionary Register access the Enumerate Register(s) using the Registry interface. They can select attributes from the Register based on their Application Schema, and bind them to the appropriate features/information/complex attributes within their Feature Catalogue, selecting only those required values (codes) from the allowable full list to satisfy the requirements for their Product Specification.
- As required, Submitting Organization representatives for a Domain can submit a proposal to the Register (and the Concept Register as required) to add new Enumerated or Codelist type attributes (and ; or new enumerated values to existing attributes within the Register. The management of content and administration of the Register is as for the Concept Register.
- NOTE possible partitioning of this Register into 2 "Domains" – a "conventional" Domain in which the attribute and its values are derived from the Concept Register; and a "classification code" Domain where the attribute is drawn from the Concept Register but the values are from a set (and mostly administered by external organizations however fixed) list of "codes" (refer to draft Register Guidelines and Conventions document, and also the first bullet above).
    - To take this a step further, perhaps the Enumerate Register can be structured so that enumerate values can be taken from the Concept Register if they are actually concepts; or registered directly in the Enumerate Register as (for want of better words) "characterizations" or "states" of a concept, in addition to classification codes. Will need to investigate this further with structures of other Registries.
- IHO GI Registry Process:
    - Data modellers working within their Domain, when requiring a registered item in the Concept Register to be an enumerate type attribute within their data model, query the Enumerate Register(s) for the existence of the attribute. If the attribute does not exist, they submit a proposal to the Register in the same manner as would be done for the Concept Register, noting however the additional hierarchical requirement to propose both the attribute and its values. Proposals are assessed by the Register Manager and Domain Control Body and actioned accordingly. The same process is followed if there is a requirement to add a new value (code) to the enumerate list for an already existing attribute.
    - Utilizing the Feature Catalogue Builder, the attribute is imported from the Enumerate Register to the Domain within the Feature Data Dictionary Register, along with those values (codes) for the attribute required for the Product Specification (which may or may not be the entire list of allowable values), and bound to features or complex attributes as required in accordance with the Application Schema.

Commented [TS7]: RM : See the previous comment.
JW: Need use cases and examples to demonstrate why the Register would need to be partitioned.
RM: I'd like a clarification of the paragraph in question later, but pending that :
Some concepts are very general and very broad, e.g., « status » in ENC vs. AIS vs. Lighthouse authority databases. The « berthing » term in the comment above.

Commented [TS8]: RM : This sounds extremely problematic. How many literals will we end up with?
JW: Agree.

Commented [TS9]: RM: Amounts to defining a derived datatype.
JW: Not sure about this comment – need to be discussed. Is there any problem with this concept?
RM: No problem. It ties into earlier comments about needing datatypes and more types of relationships between terms.

Commented [TS10]: RM : See previous comment

Product Specification Register:

- The Product Specification Register holds the published versions of all S-100 based Product Specifications.
- IHO GI Registry Process:
    - When all requirements (testing, approvals, …) for the development of the Product Specification have been met, and the final components of the specification have been produced (Product Specification document (including Data Classification and Encoding Guide); Feature Catalogue; Portrayal Catalogue), the new published version of the Product Specification is added to the Product Specification Register;
    - The previous version of the Product Specification should normally be given the status of "Retired" (note however the occasional exception as with S-52 PL, S-64 and S-58).

Submitting Organization:

- Proposals from Submitting Organizations are submitted only to the Concept or Enumerate/Codelist Registers. There is no requirement to describe how a new concept proposed to the Concept Register will be modelled in an S-100 based Product Specification – this is at the discretion of the Working Group/User Community developing their model in the Data Dictionary Register, once the concept has been registered in the Concept or Enumerate Register.
- In general, there should be at least one member of each of the Domains in the Data Dictionary Register nominated to be a Submitting Organization representative for the relevant Working Group or User Community. However, where two or more Domains are under the management of a single User Community, a single Submitting Organization representative may be identified to cover all relevant Domains (refer to IALA S-201 and S-202 Domains in the diagram); this is at the discretion of the Working Group/User Community.

Concept Register Domain Control Body (DCB):

- The Concept Register Domain Control Body is comprised of a member of each of the Domains in the Feature Data Dictionary Register.
- A member of the Domain Control Body is effectively the representative of the IHO Working Group or User Community utilizing the IHO GI Registry to develop and maintain S-100 based Product Specifications. The method by which each Domain Control Body member disseminates/discusses proposals within their expert group(s) (if at all) is at the discretion of the individual IHO Working Group or User Community for which the Domain has been created, and is therefore independent of the overall IHO GI Registry management process.
    - Example: For the S-101 ENC Domain of the FDD Register, a single person from the S-101PT (or ENCWG once S-101 is published) is appointed as the Concept Register DCB representative. When proposals are submitted to the Concept Register, the S-101 DCB representative assesses each proposal on its merits as to action to take – this may range from accepting the proposal without consultation (if for instance there is no impact on S-101); to initiating a full consultative process within the S-101PT/ENCWG (if for instance there may potentially be significant impact on S-101).