

## Paper for Consideration by Data Quality Working Group

### Input on Development of a Minimum Standard for Data Validation

**Submitted by:** Canadian Hydrographic Service and Canadian Coast Guard

**Executive Summary:** Recommendations towards the development of a minimum standard for data validation, including targeting the guideline for inclusion in S-100 to make it a common guideline for all product specifications utilizing S-100.

**Related Documents:** S-57, S-58, S-100

#### Introduction/Background

S-58 specifies the recommended validation checks for S-57 ENC, and was released May 2017 in its 6<sup>th</sup> edition. The document represents a collection of checks that have been developed and refined over nearly two decades. The document has been used by several software companies to develop validation software that is used by most hydrographic offices, RENCs and others around the world today. With the ascent of S-100 and S-100 based product specifications developed by IHO, IALA and others, it is anticipated that validation checks will also be needed for new products compiled to those specifications.

The S-101 Project Team (S-101PT) is developing a validation annex for S-101 ENC. The S-101PT is reviewing if these checks should be an annex for S-58, an annex to the S-101 Product Specification or another solution.

IALA is developing S-201, which includes over 100 validation checks adapted from S-58 due to the significant influence of S-57 on the Marine Aids to Navigation data model in the S-201.

NIPWG has developed S-122 to include validation checks, as well as S-123 which includes validation checks. In the pipeline is S-127 which will also include validation checks. All these have been developed using S-201 and S-58 as inspiration.

#### Analysis/Discussion

S-58 validation checks are divided into five sections;

1. Checks Relating to S-57 Data Structure
2. Checks Relating to the ENC Product Specification
3. Exchange Set Level Checks
4. Checks Relating to the Use of the Object Catalogue for ENC
5. Checks Relating to Allowable Attribute Values for Particular Feature Object Classes

Section 1 relates mainly to encoding (format), section 2 is a mixture between checks related to data delivery and data content, while section 3 is about data delivery. Sections 4 and 5 relate to data content. As noted these checks are dividable into three categories;

- I. Format specific validation checks

- II. Data content (data model specific) checks
- III. Data delivery checks.

The validation checks in S-58, S-201, S-122, and S-123 are all more or less individual to each product specification, but experience by the authors of this paper, and by review of various specifications currently in development tend to show that some of the checks can be fairly easily adapted to the various specifications due to general wording and application of the check.

Some additional minimum checks within the three noted categories may be possible when considering that S-100 currently has 3 encodings, two main methods of data modelling; feature based and coverage based, and one general data delivery method (exchange set). There can be similarities between these groups that can be noted and reused whenever a specification is created.

### **Recommendations**

Keep the guidance on S-100 based specification validation checks at a high level requiring S-100 specifications to include sufficient checks in the three noted categories.

Minimum sets of encoding specific checks be evaluated for each of the S-100 adopted encodings.

Minimum sets of data model checks be created for feature based data and coverage based data.

Develop the Guideline as an addition to S-100, either as a new part or as an annex to Part 4, so that it becomes common practice throughout e-Navigation development.

### **Action requested**

- I. DQWG to note this paper.
- II. Develop the Guideline on S-100 minimum set of validation check in accordance with the recommendations in this paper.