SNPWG14-11B

Proposal for Nautical Publications Data Quality

Version 0.1.1

19 January 2012

(Adapts the DQWG draft proposal to TSMAD 23 (for S-101 data quality) to Nautical Publications data quality. The text and features/attributes are light adaptations of the DQWG proposal. SNPWG should probably wait to see what is done with the DQWG proposal before taking this too far.)

1 Introduction

Data quality comprises the following:

- completeness of data (e.g. seafloor coverage).
- currency of data (e.g. temporal degradation);
- uncertainty of data;
- source of data;

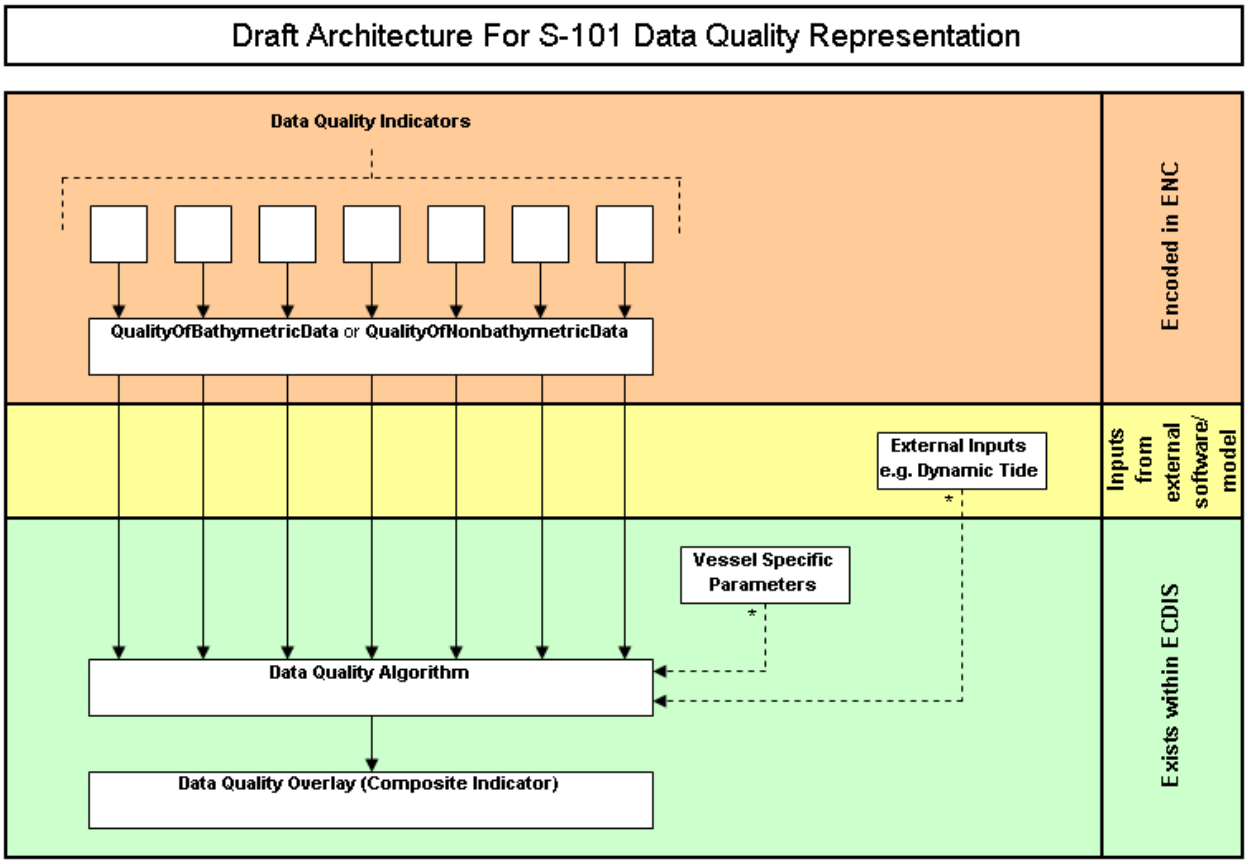
Data quality is considered to be meta information. As such, it can be encoded at three different levels (dataset, meta feature area, feature instance). All positional (2D), vertical (1D), horizontal distance (1D) and orientation (1D) uncertainty attributes concern the 95% confidence level of the variation associated with all sources of measurement, processing and visualization error. Uncertainty due to temporal variation should not be included in these attributes.

Data quality is broken into three main meta features; **QualityOfInformation**, **QualityOfBathymetricData**, **QualityOfNonbathymetricData**. The meta features cover areas where uniform assessments of data quality exist. Their scope can be limited to specific areas by limiting the area boundaries, and to specific feature classes such as “all waterways” or “all MarineServices” by specifying a feature class (or classes) in an attribute. The use of meta objects for coverage is necessary to properly express data quality for information types, bathymetry items as opposed to non-bathymetry items. A fourth class SpatialQuality is used for expressing the quality which differs from the meta objects.

Figure 1 shows the high level architecture for the revised data quality representation system used in S-101 (ENC). The individual data quality indicators (meta features and attributes) that are encoded in the ENC provide individual inputs into the data quality algorithm, which resides within the ECDIS system. This algorithm has the capability to accept additional optional inputs from vessel specific parameters (entered into the ECDIS) and external information (e.g. Dynamic tides). This algorithm then drives an on-demand data quality overlay that exists within the ECDIS system.

Nautical Publications uses the same architecture with the following changes:

- QualityOfInformation is added as a meta-feature.
- Individual data quality indicators are encoded in the NPUBS data set instead of the ENC.
- The external inputs may be different or non-existent.
- The indicators, algorithm, and overlay are those specified for NPUBS.



The NPUBS data quality model retains much of the S-101 data quality proposal, specifically classes **QualityOfBathymetricData** and **SpatialQuality**, because the NPUBS domain also includes several geographic features.

1.1 Quality of Information

The meta feature **QualityOfInformation** defines areas where uniform assessment exists for the information objects which are associated with the geographic objects in the area. It provides an assessment of the overall quality of the information objects, such as the age of the information.

1.2 Quality of Bathymetric Data

The meta feature **QualityOfBathymetricData** defines areas within which uniform assessment exists for the quality of bathymetric data, and must be used to provide an assessment of the overall quality of bathymetric data to the mariner. Areas of a dataset containing bathymetry must be covered by one or more **QualityOfBathymetricData**.

In the context of bathymetry, a feature is any object, whether manmade or not, projecting above the sea floor, which may be a danger for surface navigation. Refer to S-44.

General positional uncertainty of the bathymetry is given in **positionalUncertainty**, similarly general vertical uncertainty is given in **verticalUncertainty**. Positional or vertical uncertainty for a feature instance, that is different from the general uncertainty, can be given by using **positionalUncertainty** or **verticalUncertainty** in **SpatialQuality**.

1.3 Quality of non-bathymetric data

General positional uncertainty of the non-bathymetry features is given in positionalUncertainty, similarly general vertical uncertainty is given in verticalUncertainty. Positional uncertainty for a feature instance, that is different from the general uncertainty, can be given by using positionalUncertainty in SpatialQuality. verticalUncertainty in SpatialQuality must not be used.

Some feature types carry elevation attributes (e.g. verticalClearance). For these feature types, verticalUncertainty is an attribute on the feature instance level.

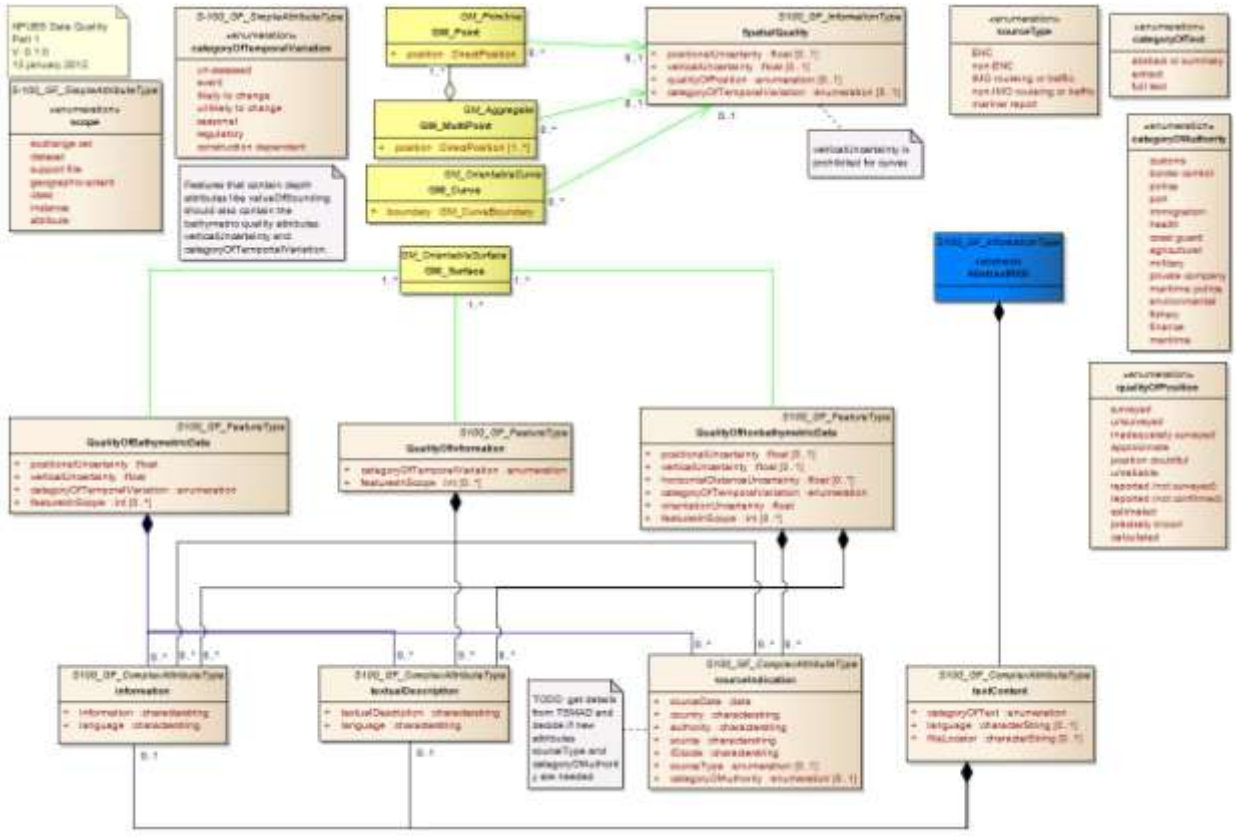
General horizontal distance and orientation uncertainty are given in horizontalDistanceUncertainty and orientationUncertainty respectively.

Some feature types carry horizontal distance attributes (e.g. horizontalClearance). For these feature types, horizontalDistanceUncertainty is an attribute on the feature instance level.

Some feature types carry orientation attributes (e.g. orientation). For these feature types, orientationUncertainty is an attribute on the feature instance level.

2 UML Model

The UML model of NPUBS data quality is given in the figure below.



3 Proposals for additions or changes to features and attributes

This section adapts the proposals in the DQWG proposal to TSMAD 23 (in TSMAD23-4.5.13B). Features defined by DQWG which are not relevant to NPUBS have been deleted. For features which are retained, attributes which have been struck through are defined by DQWG but may not be relevant to NPUBS. Attributes in *italics* have been added for NPUBS.

Proposal for a new feature type: Quality Of Bathymetric Data

Register: Hydro FDD

Name: Quality Of Bathymetric Data

camelCase: QualityOfBathymetricData

Alphacode: M_QOBD

Attribute	Type	Cardinality		Infinity	Sequential
		Lower	Upper		
~~featuresDetected~~	~~Complex~~	~~1~~	~~1~~	~~No~~	~~No~~
~~deathRangeValue1~~	~~Integer~~	~~0~~	~~1~~	~~No~~	~~No~~
~~deathRangeValue2~~	~~Integer~~	~~0~~	~~1~~	~~No~~	~~No~~
positionalUncertainty	Real	1	1	No	No
verticalUncertainty	Real	1	1	No	No
~~surveyDateEnd~~	~~Date~~	~~1~~	~~1~~	~~No~~	~~No~~
~~surveyDateStart~~	~~Date~~	~~1~~	~~1~~	~~No~~	~~No~~
~~techniqueOfVerticalMeasurement~~	~~Enumerate~~	~~1~~	~~1~~	~~No~~	~~No~~
~~fullSeafloorCoverageAchieved~~	~~Boolean~~	~~1~~	~~1~~	~~No~~	~~No~~
categoryOfTemporalVariation	Enumerate	1	1	No	No
information	Complex	0	*	Yes	No
textualDescription	Complex	0	*	Yes	No
sourceIndication	Complex	0	*	Yes	No
featuresInScope	*Integer*	*0*	*		*No*

Definition

An area within which a uniform assessment of the quality of the bathymetric data exists.

Remarks:

Nil

Proposed Change:

Add this new feature type and retire M_QUAL which this proposal will supersede.

Justification:

M_QUAL definition alludes to an all-encompassing data quality meta feature, but its use is virtually limited to bathymetric data. The new feature is clearer in its use. Furthermore the new feature type is constructed in a way that enables the retirement of CATZOC as a single ambiguous attribute.

Proposal for a new simple attribute type: Positional Uncertainty

Register: Hydro FDD

Name: Positional Uncertainty

camelCase: positionalUncertainty

Alphacode: POSUNC

Attribute type: Real

Definition

The best estimate of the uncertainty of a position in the horizontal plane.

Remarks:

Positional uncertainty is distinguished from other types of horizontal uncertainty, such as horizontal distance uncertainty and orientation uncertainty.

Proposed Change:

Add this new simple attribute type to supersede POSACC

Justification:

Change of name and definition to express uncertainty as opposed to accuracy.

Proposal for a new simple attribute type: Vertical Uncertainty

Register: Hydro FDD

Name: Vertical Uncertainty

camelCase: verticalUncertainty

Alphacode: VERUNC

Attribute type: Real

Definition

The best estimate of the uncertainty of a vertical measurement.

Remarks:

Nil

Proposed Change:

Add this new simple attribute type to supersede SOUACC

Justification:

Change of name and definition to express uncertainty as opposed to the accuracy, and to allow expression of uncertainty of all vertical measurements as opposed to only for soundings.

Proposal for a new simple attribute type: Category Of Temporal Variation

Register: Hydro FDD

Name: Category Of Temporal Variation

camelCase: categoryOfTemporalVariation

Alphacode: CATTVA

Attribute type: Enumeration

Definition

An assessment of the likelihood of change within an area since last survey	
Enumeration	Definition
Un-assessed	Temporal variation not assessed.
Event	No new survey conducted after an event (e.g. hurricane, earthquake, volcanic eruption, landslide, etc), which is considered likely to have changed the seafloor significantly.
Likely to change	Continuous or frequent change (e.g. river siltation, sand waves, seasonal storms, ice bergs, etc).
Unlikely to change	Significant change to the seafloor is not expected.
seasonal	*Varies with the season*
regulatory	*May be changed by local or other regulations*
construction dependent	*Changes likely due to anticipated or in-progress construction*

Remarks:

Nil

Proposed Change:

Add this new simple attribute type

Justification:

The attribute is needed to reduce the need for caution area or use of text in information or textual description attributes regarding the expected change in an area. Furthermore, this attribute enables data producers to express their uniform expectation of change in an area.

Proposal for a new simple attribute type: Quality Of Vertical Measurement

Register: Hydro FDD

Name: Quality Of Vertical Measurement

camelCase: qualityOfVerticalMeasurement

Alphacode: QUAVEM

Attribute type: Enumerate

Definition

Quality of a vertical measurement (Enumerates are copied from QUASOU as as suggested by DQWG.)	
depth known	the depth from chart datum to the bottom is a known value.
depth unknown	the depth from chart datum to the bottom is unknown.
doubtful sounding	a depth that may be less than indicated
unreliable sounding	a depth that is considered to be an unreliable value.
no bottom found at value shown	upon investigation the bottom was not found at this depth.
least depth known	the shoalest depth over a feature is a known value.
least depth unkown, safe clearance at depth shown	the least depth over a feature is unknown, but there is considered to be safe clearance at this depth.
value reported (not surveyed)	depth value obtained from a report but not fully surveyed
value reported (not confirmed)	depth value obtained from a report, which it has not been possible to confirm.
maintained depth	the depth at which a channel is kept by human influence, usually by dredging.
not regularly maintained	depths may be altered by human influence but will not be routinely maintained

Remarks:

Nil

Proposed Change:

Add this new simple attribute type to supersede QUASOU. List of QUASOU enumerates to be copied for qualityOfVerticalMeasurement, potentially with changes.

Justification:

Change of name and definition to extend the usefulness of the attribute to all vertical measurements, as opposed to only sounding measurements.

Proposal for a new simple attribute type: Source Type

Register: NPUBS FDD

Name: Source Type

camelCase: sourceType

Alphacode: SRCTYP

Attribute type: Enumerate

Definition

Type of source for information in the NPUBS domain	
ENC	*Derived from ENC data*
non-ENC	*Information which would otherwise be derived from an ENC but isn't for some reason*
IMO routeing or traffic	*IMO-announced routeing measure or traffic separation scheme*
non-IMO routeing or traffic	*Routeing or traffic measure not announced by the IMO*
mariner report	*Mariner's report*

Remarks:

Nil

Proposed Change:

Add this new simple attribute type

Justification:

New data quality metadata

Proposal for a new simple attribute type: Features In Scope

Register: NPUBS FDD

Name: Features In Scope

camelCase: featuresInScope

Alphacode: FEASCO

Attribute type: Integer

Definition

Feature class to which a data quality meta feature is applicable, denoted by the numeric feature codes from the feature catalogue

Remarks:

none

Proposed Change:

Add this new simple attribute type

Justification:

If data quality metadata is to be applied to specified feature classes in an area or dataset, such as "all bridges", there should be some way of specifying the feature classes.

Proposal for a new feature type: Quality Of Nonbathymetric Data

Register: Hydro FDD

Name: Quality Of Nonbathymetric Data

camelCase: QualityOfNonbathymetricData

Alphacode: M_QOND

Attribute	Type	Cardinality		Infinity	Sequential
		Lower	Upper		
positionalUncertainty	Real	0	1	No	No
verticalUncertainty	Real	0	1	No	No
~~surveyDateEnd~~	~~Date~~	~~0~~	~~1~~	~~No~~	~~No~~
~~surveyDateStart~~	~~Date~~	~~0~~	~~1~~	~~No~~	~~No~~
horizontalDistanceUncertainty	Real	0	1	No	No
categoryOfTemporalVariation	Enumerate	1	1	No	No
orientationalUncertainty	Real	0	1	No	No
information	Complex	0	*	Yes	No
textualDescription	Complex	0	*	Yes	No
sourceIndication	Complex	0	*	Yes	No
featuresInScope	*Integer*	*0*	*	*?*	*No*

Definition

An area within which a uniform assessment of the quality of the nonbathymetric data exists.

Remarks:

Nil

Proposed Change:

Add this new feature type

Justification:

QualityOfNonbathymetricData fulfill a gap in the ability to provide a uniform assessment of data quality for nonbathymetric features.

Proposal for a new simple attribute type: Horizontal Distance Uncertainty

Register: Hydro FDD

Name: Horizontal Distance Uncertainty

camelCase: horizontalDistanceUncertainty

Alphacode: HRDUNC

Attribute type: Real

Definition

The best estimate of the horizontal uncertainty of horizontal clearance and distances.

Remarks:

Nil

Proposed Change:

Add this new simple attribute type to supersede HORACC.

Justification:

Change of name and definition to express uncertainty as opposed to the accuracy.

Proposal for a new simple attribute type: Orientation Uncertainty

Register: Hydro FDD

Name: Orientation Uncertainty

camelCase: OrientationUncertainty

Alphacode: ORIUNC

Attribute type: Real

Definition

The best estimate of the uncertainty in angular measurements made relative to true north.

Remarks:

Nil

Proposed Change:

Add this new simple attribute type.

Justification:

Attribute needed to enable expression of uncertainty in angular measurements.

Proposal for a new complex attribute: Text Content

Register: NPUBS

Name: Text Content

Alpha code: TXTCON

Camel case: textContent

Attribute type: Complex

SubAttributes:

Name	***AlphaCode***	***CamelCase***	***Multiplicity***	***Sequential***
Category of text	*CATTXT*	*categoryOfText*	*1*	*n/a*
Language	*LANGGE*	*language*	*0..1*	*n/a*
Textual description	*TXTDSC*	*textualDescription*	*0..1*	*n/a*
File locator	*FILLOC*	*fileLocator*	*0..1*	*n/a*
Information	*INFORM*	*information*	*0..1*	*n/a*

Remarks:

Nil.

Conditions:

1. Either TXTDSC or INFORM must be completed in one instance of TXTCON, but not both.
2. If REGLTS/TXTCON is populated, there cannot be TXTDSC and INFORM attributes directly bound to the REGLTS.
Distinction:

Nil.

Justification:

Provides metadata to indicate to systems and users the nature, location and language of information held in external files.

Proposal for a new attribute: Category of text

Register: NPUBS

Name: Category of text

Acronym: CATTXT

Camel case: categoryOfText

Attribute type: Simple

Data type: Enumeration

Expected input:

ID Meaning INT 1 M-4

1: Abstract or summary

2: An excerpt or excerpts from a text [Should be "Extract"?]

3: Full text

Definitions:

Abstract or summary

A statement summarizing the important points of a text.

[Examples include: Anchoring forbidden; Pilot mandatory; Maximum Speed 9 knots; Wait until Green light exhibited; Under keel allowance 2.0 m.]

Extract

An excerpt or excerpts from a text.

Full text

The whole text

Remarks:

Distinction:

Justification:
Some hydrographic offices feel obliged to make the full text of a regulation available within a nautical publications dataset, because given the variety of special conditions and requirements in different countries, it is difficult or impossible to encode all the nuances of regulations using pre-defined allowed values. Others desire to encode summaries or brief extracts, because large blocks of text are more difficult to read and an important aspect of information in regulations is the conditions under which they apply and the corresponding required actions or constraints. Further, a dataset may be viewed under different circumstances (such as navigation vs. planning modes), for which varried levels of detail are

required. This attribute allows producers to designate the level of detail of the encoded text, so that multiple levels of detail can be made available in the same dataset and it is possible for applications to distinguish between levels of detail when appropriate.

Comments:

No comments.

Proposal for a new attribute: File locator

Register: NPUBS

Name: File locator

Acronym: FILLOC

Camel case: fileLocator

Attribute type: S

Definitions:

The string encodes the location of a fragment of text or other information in a support file.

Remarks:

Product specifications must state how the association between a FILLOC and a support file is made. For example, the associated support file may be identified as the file named by a TXTDSC attribute bound to the same TXTCON attribute. Product specifications are also expected to describe the meaning of the pointer for a data product and state how the supporting files must be structured (bookmarked, tagged, etc.). For example, the pointer may be an XML ID, HTML ID, line number, bookmark in a PDF file, a "key" in a resources file, etc.

Distinction:

Nil.

Justification:
This attribute simplifies the management of supporting files. It allows encoders to include in a dataset a pointer to a specific section, paragraph, segment, or other location in a supporting file. This makes it possible to have (for example) all the shipping regulations of a country in one file (or a few files) and to have different Regulations objects point to different clauses in the file.

Comment:

No comments.

Proposal for a new information type: Spatial Quality

Register: Hydro FDD

Name: Spatial Quality

camelCase: SpatialQuality

Alphacode: ?

Attribute	Type	Cardinality		Infinity	Sequential
		Lower	Upper		
positionalUncertainty	Real	0	1	No	No
verticalUncertainty	Real	0	1	No	No
qualityOfPosition	Enumerate	0	1	No	No
categoryOfTemporalVariation	Enumerate	0	1	No	No

Definition

?

Remarks:

Nil

Proposed Change:

Add this new feature type

Justification:

?

Attribute: Quality Of Position

Register: Hydro FDD (already defined in HYDRO register)

Name: Quality of Position

camelCase: qualityOfPosition

Alphacode: QUAPOS

Attribute type: Enumeration

Definition:

The degree of reliability attributed to a position	
Enumeration	Definition
surveyed	the position(s) was(were) determined by the operation of making measurements for determining the relative position of points on, above or beneath the earth's surface. Survey implies a regular, controlled survey of any date.
unsurveyed	survey data is does not exist or is very poor.
inadequately surveyed	position data is of very poor quality
approximate	a position that is considered to be less than third-order accuracy, but is generally considered to be within 30.5 metres of its correct geographic location. Also may apply to an object whose position does not remain fixed.
position doubtful	an object whose position has been reported but which is considered to be doubtful.
unreliable	an object's position obtained from questionable or unreliable data.
reported (not surveyed)	an object whose position has been reported and its position confirmed by some means other than a formal survey such as an independent report of the same object.
reported (not confirmed)	an object whose position has been reported and its position has not been confirmed.
estimated	the most probable position of an object determined from incomplete data or data of questionable accuracy.
precisely known	a position that is of a known value, such as the position of an anchor berth or other defined object.
calculated	a position that is computed from data.

Remarks:

Nil

Proposed Change:

nil

Justification:

(Evaluate whether this attribute is useful for publications data quality)