1. Introduction

1.1 The Need for Crowdsourced Bathymetry

Seventy-one percent of the Earth's surface is covered with water, yet only ten percent of the seafloor has been surveyed by echo sounders to a resolution of one arc-minute or better.

Detailed knowledge of global bathymetry is essential for creating accurate maps of the seafloor, which can in turn provide vital insight into the behavior of global earth systems and their impact on our world. The morphology of ocean basins, undersea ridges and seamounts directly influence the flow of nutrient-laden ocean currents that nourish fisheries, sustain ecosystems, and impact weather and climate.

Pollutants and marine debris are also carried by ocean currents, and comprehensive seafloor maps could provide insight into their projected path and potential effects. In addition, seafloor characteristics influence the transmission of energy from undersea seismic events, and high-resolution bathymetric models could help to predict and mitigate tsunami, storm surge or flooding impacts in coastal areas. In shallow water or areas with dynamic coastline, timely and comprehensive knowledge of the seafloor could identify submerged hazards or highlight the need for updated nautical charts and products.

With the exception of a comparatively small percentage of systematic hydrographic surveys conducted by hydrographic offices and some academic research institutions, our current view of the shape of the ocean floor is pieced together from satellite measurements of the sea surface, global gravity models and in-transit soundings collected by a 'crowd' of mariners ranging from fisherman to the academic fleet. Coordinating these existing volunteer efforts and engaging mariners to contribute data they are already collecting has the potential to exponentially expand the amount of seafloor data available. This coordination requires an approach that is both simple and easy, with just enough guidance on best practices to ensure that the resulting data is as useful as possible to the broadest range of stakeholders.

The urgent need for comprehensive bathymetric coverage will not be met by government and hydrographic office efforts alone. If we harness the collective reach of private and commercial vessels and empower mariners to 'map the gaps' in seafloor data in a coordinated way, we can greatly increase our understanding of the seafloor, and its influence on the world around us.

1.2 Purpose and Scope of this Document

The purpose of this document is to provide guidance to mariners to help them collect and contribute bathymetric data in a format that will allow the data to be useful to the broadest possible audience. It is hoped that this document will help mariners to leverage data that is already being collected on vessels equipped with common commercial echo sounders and Global Navigation Satellite System (GNSS) receivers, and will provide them with information about devices, techniques and formats that are approved by the International Hydrographic Organization for crowdsourced bathymetry (CSB) data sharing.

This document also provides some information about position and depth accuracy and data uncertainty, to help the mariner better understand some of the considerations and limitations of crowdsourced bathymetry, as well as the feasibility of using the data for certain applications. The legal considerations of bathymetric data logging and crowdsourced bathymetry sharing are also briefly explored.

This document does not provide guidance on how best to use crowdsourced data as an end customer, but it does provide links to existing documents and sites that already address that topic. The important social engagement aspects of crowdsourcing such as gamification and recognition are also beyond the scope of this document, although it is acknowledged that the scope of CSB is far-reaching and has many potential future applications.

1.3 Target Audience

First and foremost, this text seeks to inform and guide potential individual collectors and contributors of crowdsourced bathymetry data. In addition, organizations (also referred to as 'trusted nodes') interested in serving as liaisons between contributing mariners and the International Hydrographic Organization (IHO) should find this information helpful. Users of crowdsourced bathymetry data may also find this document informative, although they are not the primary audience.

1.4 Document Structure

This document addresses seven topics related to crowdsourced bathymetry. The first topic, "Overview of Systems and Sensors," provides basic information about systems, sensors and concepts that are necessary for collecting bathymetric data. The second topic, "Metadata," details a standard metadata structure for crowdsourced bathymetry datasets which will facilitate easy exchange of data. The next topic, "Data Collection," outlines hardware and software considerations for logging CSB information. The section on "Uncertainty" delves into data quality issues, and discusses how mariners and end users can better understand the impact of various factors on the reliability of a dataset. The section on "Data Contribution" provides guidelines for submitting CSB to the IHO, and that section is followed by a description of the “IHO’s Data Centre for Digital Bathymetry,” which will warehouse crowdsourced data contributions. The final section, “Legal Issues,” discusses the legality of collecting and sharing data.

The guidance in each section is provided by subject-matter experts and is informed by ongoing crowdsourced bathymetry pilot projects. Additional detail and further reading are provided via Annexes and external links. This text is intended to be a living document, and will be updated as improvements are identified, and with feedback from the marine community.

2 Overview of Systems and Sensors

2.1 Echo Sounders

Echo sounders determine water depth by transmitting sound pulses from a transducer through the water to the seafloor. The time interval between the emission and return of a pulse is recorded, and the distance from the transducer to the seafloor is determined by multiplying half the time interval to the speed of sound in salt water.

The speed of a sound wave traveling through water is influenced by the temperature, salinity, and depth of the water column. For most mariners, setting an echo sounder to an average sound velocity constant of 1500 m/s (in saltwater) is adequate for general navigation. When a very precise understanding of seafloor depths is desirable (such as for coastal hydrographic surveys) scientists may deploy sensors to directly measure the composition of the water column. These measurements are then used to calculate sound speed variations and generate a very accurate determination of water depths.

[Insert diagram of how echo sounders work, and how different chemical and physical environments impact data quality]

Types of echo sounders and their technology are described in Appendix A.

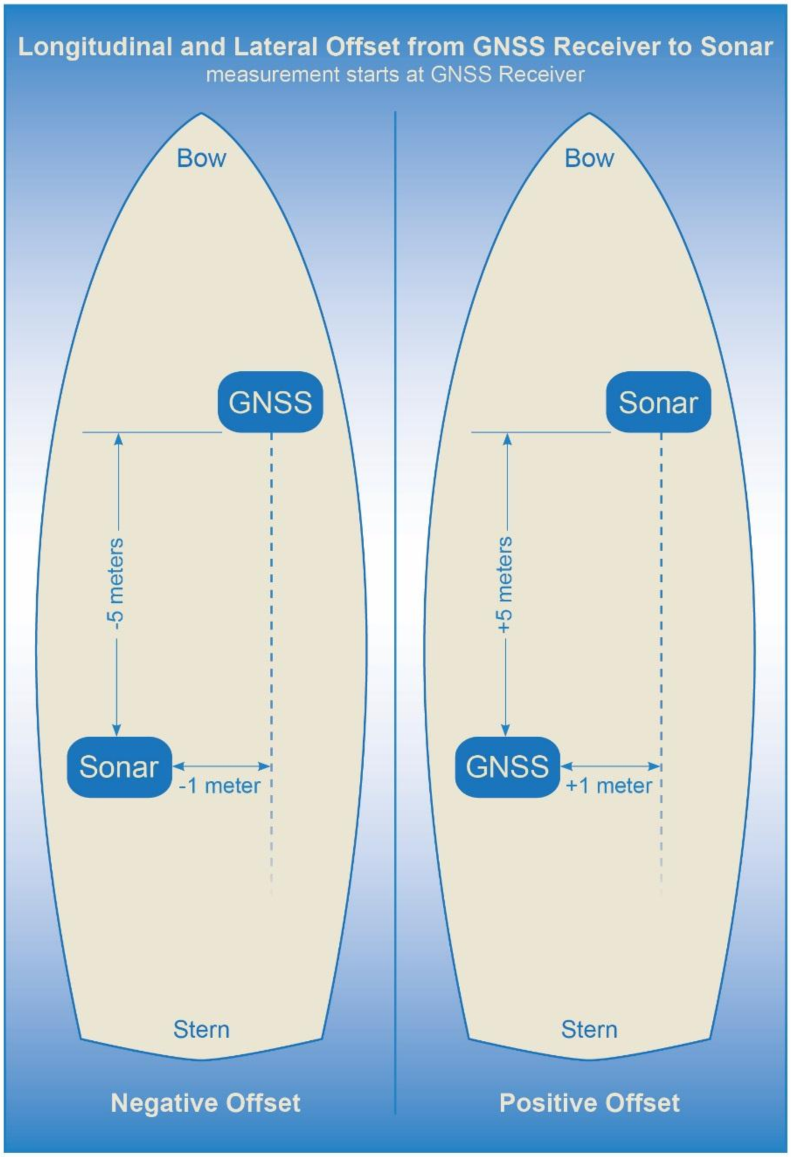
2.2 Positioning Systems

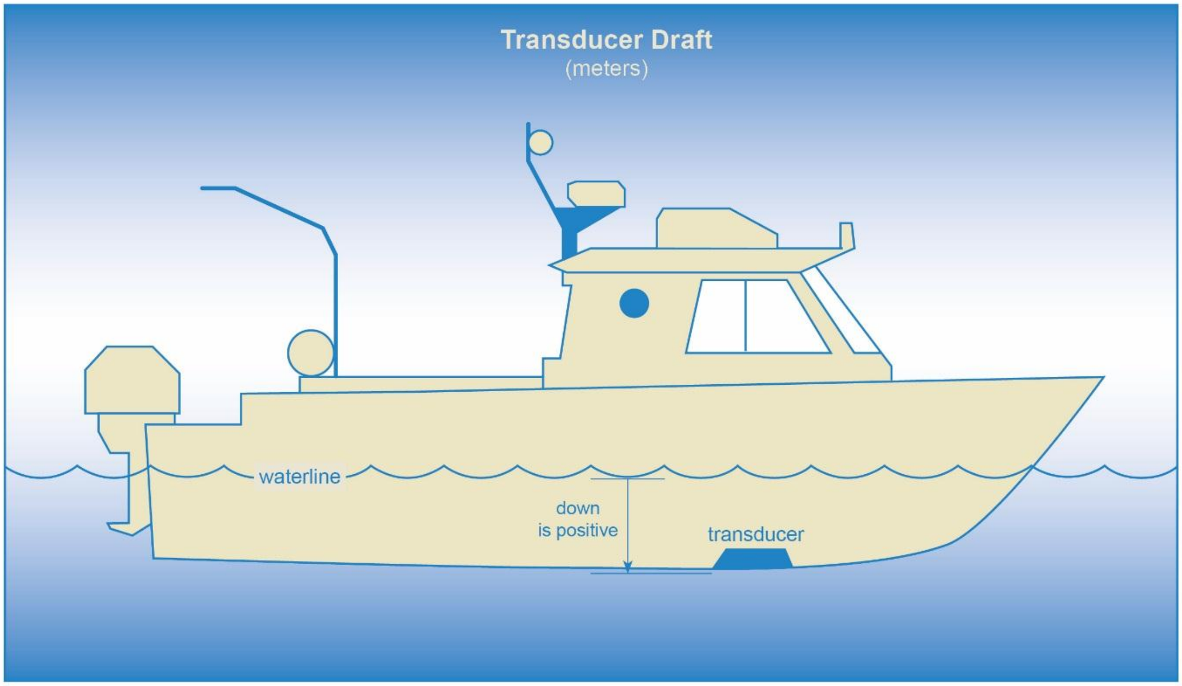
2.3 Motion Sensors

2.4 Data Loggers

2.5 Measuring Offsets

[Include definition, demonstrate impact on final data and quality


Longitudinal and Lateral Offset from GNSS Receiver to Sonar
measurement starts at GNSS Receiver
Bow
GNSS
-5 meters
Sonar
-1 meter
Stern
Negative Offset
Bow
Sonar
+5 meters
GNSS
+1 meter
Stern
Positive Offset



2.6 Integrated Systems

[Suggest diagram of integrated system, i.e. echo sounder, positional system, motion sensor, and data logger – and the data flow within the system.]

3 Metadata

3.1 Introduction to Metadata

3.1.1 What is Metadata?

It is important to understand the difference between data and metadata. Simply put, data is information, and metadata is information about the data. In the context of crowdsourced bathymetry, the data are the individual point observations from the vessel (consisting of a depth, date, time, and geographic position), whereas the metadata provides the user with additional facts about the dataset. For example, metadata can provide information about where the echo sounder's transducer is mounted on a ship's keel, and what type of vessel collected the data.

3.1.2 Why is Metadata Necessary?

Crowdsourcing, by definition, is the collection of data from a large number of untrained observers. While today's professional mariners are hardly what one would call 'untrained,' depth measurements from the 'crowd' are collected using a wide variety of sensors and techniques. While some variability is to be expected, for data to be useful and interpretable there must be some unifying information that describes the factors influencing data collection. Metadata provides that information, and ensures that bathymetric data have consistent collection criteria, nomenclature, and structure.

The metadata associated with crowdsourced bathymetry is vital, because it allows future end-users to make informed decisions about the quality and potential applications of the dataset and to apply enhancements or corrections if necessary. For example, documentation about the time and date when a depth measurement is taken allows a user to apply tidal corrections to data. Similarly, information about a transducer's vertical offset from the waterline, or its horizontal offset from a GPS receiver, allows a user to apply vessel draft and horizontal positioning corrections to the data. By applying data corrections based on information in the metadata, end-users are able to improve the value of the bathymetric data.

In addition to informing a user about a dataset, a uniform metadata structure ensures that data is consistent and accessible, regardless of the platform that collected the information. The subsequent section provides guidelines for data and metadata structures that will allow mariners to collect data in a format that is uniform and allows for easy data-sharing and crowdsourcing.

3.2 Data and Metadata Descriptions

[Add text that describes unique identifier requirement]
These tables conform to the CSB GeoJSON data format 2.0 [include hyperlink to GeoJSON]

Table 3.1

Data Field	Description	Example
Longitude *(Required)	Describes the longitude value of the horizontal geographic position; in WGS84; Decimal Degrees to six decimal places; Normally derived from the GPS NMEA GGA String; Positive = East; Negative = West	19.005236
Latitude *(Required)	Describes the latitude value of the horizontal geographic position; in WGS84; Decimal Degrees to six decimal places; Normally derived from the GPS NMEA GGA String; Positive = North; Negative = South (request as many decimals as can be provided)	40.914812
Depth *(Required)	Describes the measured distance to the sea floor. Depth is always a positive value in *meters with accuracy of tenths of meters;* Normally derived from NMEA DPT data string	7.3
Date & Time *(Required)	ISO 8601/UTC Time Stamp of the depth measurement; Normally derived from NMEA GGA string	2015-08-06T22:00:00Z

Table 3.2

Metadata Field	Description	Example
Convention	Describes what CSB JSON format version used	CSB 2.0
Unique Vessel ID *(Required) "platform.uniqueID"	Identifies the trusted node source, and uniquely identifies the contributing vessel; the first five characters designate the trusted node source, the sixth character is a hyphen (-), and the remaining characters are generated using a UUID. The UUID should be consistent for each contributing vessel throughout the life of service of the vessel.	SEAID-*UUID* *Use ROSEP for Rose Point Navigation contributors*
Platform Type	Describes type of vessel; Broken down to general vessel types, such as fishing, recreational, sailing, cargo, tanker,	Options for type value: Fishing, Tug, Sailing Vessel, Recreational Craft,

	tug/tow, etc… see document List of types will be provided to contributors (trusted nodes) via web API	Passenger, Cargo, Tanker, Research Vessel, Other
Platform Name	The name of the vessel; open string	White Rose of Drachs
Platform Length	Length of the vessel; a positive value in meters, with an accuracy to the nearest meter	65
Platform Length Unit of Measure	Always in meters	meters
ID Type	This designates the ID number provided. Vessels can choose only one type. Currently, only two types of vessel ID numbers are available: MMSI and IMO. If these IDs are unknown or not assigned, a None value is accepted.	Options for ID Type: IMO, MMSI, None
ID Number	Provides the input value for the ID Type chosen above	369958000
Sensor Type Sounder	Defines the sensor type for echo sounders. This must always be defined as: “Sounder” (not an optional field that users can change)	Sounder
Sounder Make	Free text. In the future, a list of sounder makes will be provided to contributors (trusted nodes) via web API.	Sperry Marine (L3 ELAC)
Sounder Model	Free text. In the future, a list of sounder models will be provided to contributors (trusted nodes) via web API.	ES155100-2
Sounder Transducer	Free text. In the future, a list of echo sounder transducer options will be provided to contributors (trusted nodes) via web API.	Dual Freq 200/400 kHz
Sensor Type GNSS	Defines the sensor type for GNSS receivers. This must always be defined as:	GNSS

	"GNSS" (not an optional field that users can change)	
GNSS Make	Free text. In the future, a list of GNSS receiver makes will be provided to contributors (trusted nodes) via web API.	Litton Marine Systems
GNSS Model	Free text. In the future, a list of GNSS receiver models will be provided to contributors (trusted nodes) via web API.	LMX420
Sounder Draft	Vertical distance in meters from the waterline to the vessel's transducer. Draft value is always a positive value in meters with accuracy of tenths of meters	4.6
Sounder Draft Unit of Measure	Always defined as meters	meters
Sounder Draft Applied	Boolean - true or false - designation for reporting whether or not the Depth values reported in the file have been vertically corrected for the Draft offset.	false
Longitudinal Offset from GNSS to Sounder	Longitudinal offset from GNSS receiver to sounder. Values are in meters, positive moving from the stern to bow; i.e. when the GNSS receiver is aft of the sounder, the value is positive; and when the GNSS receiver is forward of the sounder, the value is negative. Accuracy given to the hundredths of meters.	3.52
Longitudinal Offset Unit of Measure	Always defined as meters	meters
Lateral Offset from GNSS to Sounder	Lateral offset from GNSS receiver to sounder. Values are in meters, positive moving from port to starboard; i.e. when the GNSS receiver is on the port side of the sounder, the value is positive. Accuracy given to the hundredths of meters.	-0.76
Lateral Offset Unit of	Always defined as meters	meters

Measure		
Position Offsets Applied	Boolean - true or false - designation for reporting whether or not the final Position (Lon, Lat) reported in the file has been corrected for the lateral and longitudinal offsets between the GNSS receiver and the sounder.	false
Sound Speed	Value describes the sound speed used by the sounder to calculate the distance to the sea floor. Value is reported in meters per second (m/s). The average speed of sound in seawater is around 1500 m/s, and this is normally the default value used by echo sounder processing units. If left blank, users will assume 1500 m/s, unless they use an overriding sound speed environmental model.	1500
Sound Speed Unit of Measure	Always defined as meters per second (m/s)	m/s
Provider Contact Point Organization Name (orgName)	Trusted Node Source Name	Sea-ID
Provider Email	Trusted Node Source email address. Used as contact point when users of the data want more information. (not required)	support@sea-id.org
Provider Logger	Name of software program or hardware logger used	Rose Point ECS
Provider Logger Version	Version of the software or hardware logger used	1.0
Depth Units	Designates the depth measurement units Always defined as meters	meters
Time Units	Designates the time measurement units Always defined as ISO 8601/UTC	ISO 8601

4 Data Collection

4.1 Introduction

Bathymetric data collection from multiple sources has both advantages and disadvantages for technical requirements. This diversity of sources means there is a wide variety of software, hardware, and methods being used and explored. Vessels range in size and capability and may present unique limitations to collecting bathymetric data. Pilot projects are currently underway developing solutions to these limitations and developing systems that are as universal as possible.

For instance, recreational boaters most likely don't have a large 'below decks' area or VSAT (Very Small Aperture Terminal) internet. They will be better served with a piece of hardware that easily survives a power cycle and stores its data on an external USB stick that they can unplug when the day is over. Likewise, a superyacht or bulk carrier is better served by a totally hands-off solution that uploads its data through the VSAT internet link every day.

Through pilot projects, it is hoped that a solutions for vessels' limitations will be identified.

Diversity of sources can also create a challenge in data exchange formats and project requests and requirements. It is hoped that projects will remain actively involved with the Crowd Sourced Bathymetry Working Group (CSBWG) at the IHO which works to standardize these processes. In other words, whichever project the mariner chooses to contribute data through, if the project is involved in the CSBWG, the data will be usable.

4.2 Crowd Sourced Bathymetry Data Collection Process

There are three steps in the CSB gathering and transmittal process: 1) data collection; 2) storage; and 3) exchange. [suggest diagram(s) for this section]

4.2.1 Step One: Data Collection

Some data collection and exchange systems or "setups" require only minimal software to be active. For instance, if you have a piece of equipment on board that has the basic required inputs on it, such as an ECDIS system that has input from the depth sounder and the GPS, then the ECDIS manufacturer could easily start recording those soundings and exchange them with a data center (on land) through the system's update mechanism. When charts are updated, the system could also upload the recorded soundings to them. However, very few setups provide such an easy process.

Typically, the data collector will be connecting a separate data logger to the two minimum inputs: (1) input from the depth sounder, and (2) input from the GPS unit.

Most instruments on the bridge that have to do with navigation speak a professionally developed standard language to exchange messages, NMEA (National Marine Electronics Association). The NMEA 0183 standard describes what instruments have to adhere to if they want to exchange data using this protocol.

Most instruments have an output port that have a pair of copper wires in them on which “sentences” are broadcast at the very low speed of 4800 baud (which is slower than the 9600 baud dial-in modems from the early years of the internet). The slow speed is irrelevant given the minimal amount of that that is collected and exchanged— every few seconds the GPS will send the vessel’s position, the depth sounder will send the depth under keel, and the anemometer (i.e. wind speed meter) will send the direction and strength of the wind. The instruments provide these data or “sentences” in the NMEA format so that it is understandable for many users.

An actual NMEA sentence from a GPS unit may look like this, where latitude and longitude are visible:

$GPGLL,0424.99,N,11359.77,E,012636.21,A,D,*5E

The same GPS may also provide a sentence that looks like this:

$GPGGA,071953.00,0424.9862,N,11359.7661,E,1,9,1.8,21,M,,M,,*68

Again, the position can be derived from this sentence. It also includes information about the accuracy, altitude, and time the GPS unit acquired the position.

The same logic is used by the depth instrument. It gets input from the transducer which is mounted outside the hull. The depth instrument outputs a sentence like this:

$SDDBT,0006.0,f,0001.828,M,0001.0,F*3A

In this sentence, DBT stands for depth below transducer. The depth in feet, meters, and fathoms are also visible.

Integrated instruments on a bridge can read inputs from various sources, such as wind, water temperature, depth and position, and perhaps even info from the autopilot. These integrated instruments produce different sentences. [provide link to definition/example of sentence.]

The description for installing a data logger can be found in section [insert section or chapter number].

In order to start logging bathymetry data, instruments must connect to the data logger via the aforementioned copper cables (that carry the “sentences”), the cable(s) that carry the DBT and either the GLL [insert acronym definition] or GGA [insert acronym definition].

The GLL and GGA measurements include time, enabling the calculation of derivative information or to correct for things such as tides and swell. Storing the date and time with every depth/position pair is highly valuable (best acquired by GPS in UTC).

4.2.2 Step Two: Storing the Data

Single-beam echo sounders provide a simple, narrow depth reading straight down from the transducer to the bottom of the sea. Common single-beam echo sounders record two megabytes of data or less (less than the size of a picture taken by an iPhone). Multi-beam sonars, which are much more expensive and can be found on survey vessels and some newer professional ships (e.g. fishing vessels and super yachts) provide a continuous swath of detailed depth information beneath the ship. A multi-beam echo sounder records gigabytes of data per day.

4.2.3 Step Three: Data Exchange

After the data is collected and stored, it needs to be transmitted. Sending and receiving data at sea is more challenging than on land. It is possible to receive a 3G/4G signal on a phone or through a 3G/4G router on board. However, once the vessel is beyond sight of land, it may be necessary to subscribe to a satellite service to communicate with any party on land.

AIS (Automatic Information System), commonly used to send messages between vessels, can be used for certain bathymetric applications. However, transmitting whole data log files of depth soundings via AIS is inefficient.

Depending on which CSB project, or trusted node, you work with, their data logger will have a preferred (if not restricted) method of transferring data (see section 5.5). Some methods of data transmission are simple, such as sending a USB stick via mail to the data center, connecting to a smartphone via Bluetooth to upload files, or directly plugging into a VSAT modem. Other methods are expensive. Inmarsat SAT-C or FBB [insert acronym definitions and links] is billed per-megabyte, so transmitting the data on land via Wi-Fi may be cost-saving. CSB data typically isn’t time sensitive: as long as data is consistently sent to the datacenter, it’s a valuable contribution.

4.3 Best Practices and Recommendations

Many of the existing CSB projects have developed “best practices,” outlining what should and should not be done in order to be a successful contributor. [link to example of best practices] These might be primarily of interest to developers, but anyone interested in CSB can learn from these.

4.3.1 Keep the Data in the Original Format

Stripping data from an NMEA sentence and only saving parts of it is not recommended. Saving the data in its original format will help validate data recordings and troubleshoot potential anomalies in the data. For example, if only *depth in meters* from the DBT string is saved, then a strange reading cannot be compared to the depth in feet. Or, if only the latitude and longitude for position are saved, the detailed information in the GGA sentence such as a quality assessment and the timestamp on which the fix was taken will be missing.

While the IHO Data Center for Digital Bathymetry (DCDB) only accepts GeoJSON or XYZT data, in which the depth and position are only noted in one format, having the additional data available is highly recommended.

4.3.2 GPS Latency and Quality

As mentioned above, GGA provides more information than GLL. If both sentences are available, save GGA. If the GPS unit provides even more information (such as latency), save that also. Only a few units do.

4.3.3 Real-Time Clock

The internal clock of a computer typically runs 'off' by several seconds per week and synchronizing the clock by NTP (Network Time Protocol) is only possible if there is a network connection.

The GGA and GLL sentences from a GPS unit provide all the info needed. Set the system clock to the correct date and time (you need the date anyway) for logging and debugging, but use the time provided by the GPS and set the internal clock to that input if you can.

Some systems use internal counters (first recording, second recording, and so on) but this is not recommended. If it is necessary to rely on the system clock for the date, document (and save) the process of setting this, and investigate how it will behave after a long period without power to the system.

4.3.4 Time Synchronization of Sensor Input

The NMEA sentences, a first stream from the GPS unit and a second from the depth sounder, will come at intervals dictated by the unit's capacities. The GPS might send a location sentence every second; the depth sounder might send one every three seconds.

At its simplest form, it is necessary to 'couple' both bits of information as well as possible. At its most complex, calculate 'where precisely' the depth sounder was at the time it took its reading—meaning, calculate the location based on the timestamp of the depth measurement, and find the spot in between the two closest position measurements.

It is essential to store all measurements with an accurate timestamp 'at the time of measure' and then allow for the complex calculations to happen in post processing (even in the datacenter if the collected data is not used onboard). Saving timestamps with every reading allows data to be re-processed if you change or improve interpolation methods, even years later. This approach is strongly recommended.

4.3.5 Varying Draft, Keel Depth

As described above, the draft and position underwater of the depth sounder can be transmitted with the collected data. If the vessel collecting data has a varying draft (i.e. if it takes a lot of fuel on or offloads goods), it is important to collect this information and connect it to the series of depth readings.

It can be as simple as storing, with a timestamp, the current draft of the vessel. So, every time it changes significantly, record it in the stream of collected data. This will allow for adjustments during post-processing.

4.3.6 Compliance with NMEA Specification & Specifically with Opto-isolation

As mentioned above NMEA has been developing and promoting the standard through which, for many years, messages have been sent between bridge equipment. NMEA also makes sure that bridge hardware is developed according to the same standards. For further information, see the NMEA standard on opto-isolation. [insert link]

Hardware interoperability is critical. In short, one piece of hardware could negatively affect the other. On an integrated bridge, where everything is connected, it would be difficult to identify and locate one piece of equipment that is malfunctioning. Isolating the signal through optometry, away from the small current on a copper wire, is one way to ensure this.

It is recommended that this potential issue is at least considered on larger vessels and on certain data loggers. Vessels that can only fit type-approved hardware will request it. Hardware developers considering setting up a CSB project should keep this in mind. Mariners selecting a CSB project to work, should consult a professional bridge equipment installer on this issue, and ensure the NMEA standard is respected throughout.

4.3.7 Continuity of Electrical Power

Continuous power aboard ships is never a guarantee. Some vessel invest in a large and well-maintained Universal Power Supply (UPS) for all of the bridge equipment. However, there are still times when the transition from shore power to a generate causes a momentary loss in power and data loggers must reboot and recover. Consider investing in a small, serviceable, built-in battery to the logger to ensure smooth operation.

4.3.8 Hands-free Operation

The best results from any crowd-sourced project occur when the user is passive and measurements are purely based on the technology doing its job. For example, when Google developed an algorithm for calculating average speed on roads, they found that once they started tracking phones in cars, rather than user-reported data, the algorithms got far more accurate. Relying on a user to 'report' data isn't consistent, and supports only a certain kind of reports (such as slow-downs, accidents, etc.) not the all of the information available.

There are times, however, where the user's interaction is valuable and even required, for example, when the transducer depth or vessel draft changes or the setup is changed, such as the GPS antenna or the position of a sensor.

A data logger and its setup should allow the user to record these changes at the time they occur or after the occurrence (e.g. when they've taken fuel and are underway again, or even days later when they remember their draft has changed significantly).

5 Uncertainty

One of the most commonly asked questions about crowdsourced bathymetry is: can I trust the depth measurements? The answer to this deceptively simple question is, unfortunately, both yes and no. Crowdsourced bathymetric data are subject to a number of issues that cause them to differ from the true value of the data being measured.

For example, the estimation of depth from an echo sounder relies on a measurement of time, which is then converted to depth based on an assumption about the speed of sound in the water. If the estimate of speed of sound used is wrong, then all of the depths are also going to be incorrect to some degree. Similarly, when the echo sounder program examines the sound data returned from the water around it, it has to determine which echo represents the seafloor, and when that echo happened relative to the time of transmission of the sound. If the first part of the problem is done incorrectly (such as representing the depth of a fish in the water rather than the seafloor) or the second part is done incorrectly (such as acoustic noise in the area from other boats) then whole depth could be incorrect. The former case is often simpler to detect, since the error can be large and therefore obvious; the latter can be more subtle.

In any of these cases, the measurements are subject to a number of causes of uncertainty about the true depth in any particular area, and even the location of the measurement once the effects of GPS positioning are taken into account. It is therefore necessary to qualify these data with some estimate of the uncertainty of the data, and that we take this, and other effects, into account when the data is processed, stored, and used.

For example, if the depth is reported as 10m, is it 10m ± 0.5m[1], or 10m ± 1.0m? Alternately, is it 10m ± 0.5m with the most likely value at 10m, or is it 10m ± 0.5m with the most likely value consistently either shallower or deeper than 10m? The sources, consequences, and methods associated with the estimation and manipulation of these quantities are the subject of this section of the document.

A full treatment of all of the types of uncertainty that can affect bathymetric data would be extremely long and of interest to at most a minimal subset of the potential audience of this document. The approach taken here, therefore, is to present the features of uncertainty that are likely to be important to everyone first, and then to address the issues relevant to each of three categories of users: individual observers, trusted nodes, and end users of the resulting database. In each case, the material in the body of the document reflects the most important information and methods appropriate to the user group. Further details are presented in the supplementary material in Appendix X.

[1] This is not an entirely valid way to express uncertainty, since it does not address either the type of distribution expected, or the level of significance or coverage factor associated with the statement. These details are explained more fully following, and in Appendix X.

5.1 Meaning, Sources, and Consequences of Uncertainty

5.1.1 The Meaning of Uncertainty

The term “uncertainty” is, unfortunately, used both technically and colloquially, leading to misunderstandings as to its meaning. In a scientific context, “uncertainty” is understood to mean a measure of how significantly different a particular measurement could be from its true value. Ideally, the best way to do this would be to compare the field-acquired data against the corresponding true value. Unfortunately, however, the true value for any given measurement is usually both unknown and unknowable, and it is therefore impossible to compute the error in a measurement (i.e., the difference between the measured and true value) directly. The best we can do is to estimate the scale (in magnitude and sign) of the error that we believe may be committed, and then express that in some useful summary fashion. This summary is the uncertainty of the measurement.

It is important to distinguish in this that “uncertainty” in this context is distinct from any blunder that might be committed. (Unhelpfully, “blunders” are sometimes also called “errors” which can lead to more confusion.) A blunder is a measurement that is incorrectly made, whereas any measurement, even made as well as we possibly can, still has some degree of uncertainty.

For example, consider the case of measuring the depth of water using an echo sounder. Assume for the time being that all the appropriate corrections are being done---the echo sounder is correctly installed and calibrated, the sound speed profile is measured, and the motion of the vessel is being tracked. For any particular measurement cycle, the echo sounder has some algorithm that examines the acoustic signal being returned from the water and attempts to determine the depth of the water. No algorithm is perfect, however, and it is entirely possible that the echo sounder might mistrigger on some loud event in the water that occurred prior to the sound intersecting with the seabed: a large fish school, for example. The echo sounder therefore commits a “blunder,” misreporting the depth.

On the other hand, even assuming that the echo sounder is not fooled by any other noise in the water, the depth reported for any measurement cycle is going to be different, even if the echo sounder were strapped in place and continually observing the same seafloor. Small changes in the local environment can result in different signals being returned, which cause the bottom detection algorithm to change its results slightly, resulting in different depths being reported. This is uncertainty in the sense meant in this section, and the extent to which this uncertainty materially affects the depths being provided is the focus of this discussion.

Uncertainty, then, provides a qualification of the measurement being reported. The goal is to provide, for the end user of the data, a statement as to the potential variability, or bounds, within which the true value of a measurement is expected to lie. Doing so allows the end user to judge whether the data is suitable for the purpose intended, and allows for the comparison of data, or selection of appropriate processing techniques. In a very strong sense, measurements of estimates without an associated uncertainty, is essentially incomplete.

5.1.2 Sources of Uncertainty

All of the measurements that are made to support bathymetric mapping are heavily composite, meaning that a number of different measurements are made and then combined to construct them. In order to maintain some level of control over the complexity that can ensue, it is common to attempt to categorize the different types of uncertainties that can occur, and then estimate their magnitudes before combining them together for some specific purpose. (The methods of combination can change by purpose.) The four main sources of uncertainty that are applicable to crowd-sourced bathymetry are:

- **Stochastic (random) uncertainty**. This is the uncertainty of individual measurements which is expected to vary randomly from measurement to measurement. Because of this, this type of uncertainty can sometimes be reduced by taking multiple measurements of the same thing, and then averaging: the random variations average out, and the true measurement remains. The ping-to-ping variation in depth estimated from the face of the transducer to the seafloor is this type of uncertainty.

- **Systematic uncertainty.** This is the uncertainty associated with a correction that should in theory be possible to implement, but which in practice is not done, leaving every measurement conducted by the instrument in error by a given amount. Because the amount by which the measurement is in error is not known, an allowance has to be given in estimating the overall uncertainty of the measurement. The uncertainty caused by not measuring the sound speed in water before converting time-of-flight to depth is this type of uncertainty.

- **Integration uncertainty**. This is an uncertainty associated with failure to install an instrument correctly, or for failing to adequately document the installations that were done. This could be considered a type of systematic uncertainty, since the behaviors of the uncertainty are very similar, but is often better considered separately, since it is something that can reliably be done for all systems. The uncertainty caused by not measuring the offset of the sonar transducer from the waterline is this type of uncertainty.

- **Modeling uncertainty.** Every measurement is an abstraction of the real world, based on some assumptions that allow us to simplify the complexity of the world in order to be able to do computations. For example, we know that the seafloor has many different levels of complexity associated with it, but often assume that we can model it as a continuous mathematical surface in order to be able to interpolate a likely depth where there was no actual measurement. The degree to which the model used accurately reflects the real world can cause there to be a difference in the model's output compared to the real world. Since the alternative is not to do the computations, we often willingly accept such approximations, but include a component of uncertainty to reflect the fact. Most difficult of the uncertainties to estimate, and often ignored, the uncertainty caused by assuming a particular mathematical form for the seafloor is this type of uncertainty.

In addition, although it might be considered a form of modeling uncertainty, many datasets are incomplete, meaning that they do not contain sufficient data to completely specify the measurements being reported, or the products which are subsequently constructed. For example, if a dataset consists of measurements of depth that are more than 50m apart everywhere, it is impossible to assess the shape, location, or presence of objects smaller than approximately 100m - and often significantly larger. It is, of course, possible to interpolate the data to an arbitrary resolution - generating, say, a 1m grid. However, the information in this grid at the smaller scales is mostly an artifact of the assumptions built in to the interpolation scheme, rather than what is actually present in the real world.

Consequences of not understanding or accommodating this behavior are that significant objects in the real world might be missed, resulting in data that appear to be accurate, but which do not reflect the actual environment. Data, and in particular gridded data, can be very visually persuasive, however, which can result in pernicious, but unwarranted, belief that the data are better (i.e., of higher reliability or lower uncertainty) than they actually are. Estimating and reflecting uncertainty of this type is, however, significantly more complex than the other forms of uncertainty considered here. More details and suggested references are provided in Appendix X.

As a way of controlling the complexity of estimating the uncertainty of a measurement or system that includes a number of different components, it is often the case that the various forms of uncertainty are estimated separately, and are then combined into an overall assessment of the uncertainty. This works well when there is sufficient information available to help with the assessment of the uncertainty, typically through analysis of data, or the associated metadata. For many of the types of uncertainty, however, it can be very difficult to obtain data that allows for the uncertainty to be estimated directly.

For example, imagine the case where an echo sounder is installed below the waterline, but reports depth from the front face of the transducer to the seafloor. This would typically be considered an integration uncertainty, but also has aspects of a systematic uncertainty. Depending on the loading of the vessel, and its motion through the water (taking into account speed, heading, currents, etc.), the distance from the mean water surface to the transducer can vary significantly. If the data is not corrected for this effect, depths reported by the echo sounder have a time-, position-, and situation-varying uncertainty imposed on the primary signal that indicates true depth. Since none of the information that is required to correct for these effects is being recorded, it is essentially impossible to correct for them after the fact. Consequently, they either have to be treated as one large uncertainty that encompasses all of the uncertainty being induced, or the data have to be processed against other references in order to approximate the required corrections. This process degrades the quality of the data, or increases the cost of ownership, or both. If the (lack of) corrections are not documented, then the situation is even worse. Hence, good installation practice, and metadata provision, are essential to the overall uncertainty of the data: many effects cannot be corrected after the fact.

Which sources of uncertainty are important depends on the task to which the data is being put. For example, an individual observer might be more concerned with the measurement uncertainty of the particular echo sounder being used on their vessel, while a trusted node would be concerned about uncertainties between vessels, and a database user might be more concerned about the modeling uncertainty inherent in the modeling technique being used to develop the data from the database into a product. More details on particular uncertainties, and suggestions for best practice, are provided in the sections below focusing on these use cases.

5.1.3 Estimation & Expression of Uncertainty

The most common method for estimating and expressing uncertainty is through some statement of the statistics associated with a measurement or system. Ideally, this is done by making the same observation multiple times, and then assessing the degree to which the measurement indicated changes between different observations. For example, imagine the case where an echo sounder is mounted in a tank, and constantly measures the depth. Since we do not believe that the depth of the tank is varying significantly within the duration of the observations (e.g., due to atmospheric pressure, evaporation, or condensation), or we consider these variations to be insignificant with respect to the scale of the variations being examined, any change in the indicated depth must be due to uncontrolled variabilities in the echo sounder. For example, there could be differences due to acoustic or electrical noise in the returned signal. Taking all of the measurements together, it would be possible to estimate the average depth returned, and therefore the degree of variability of the depths about this average value. If the depth of the tank was independently measured, for example by draining all of the water and measuring with a laser level, then it would also be possible to estimate any bias between the average acoustic depth and the physically measured depth[2]. Going further, it might be possible, depending on the number of observations, to estimate whether the observations are all clustered symmetrically around the average value, or if there is an asymmetry in the observations, with most likely value being shallow or deeper than the average.

This ideal case rarely occurs in practice. In many physical systems it is very difficult to keep conditions sufficiently constant that multiple observations can be made of exactly the same system. For example, imagine that the echo sounder is now attached to a dock and observes ostensibly the same patch of seafloor on each measurement cycle. Changes in the water properties on the scale of a few measurement cycles can cause sufficient difference to dominate the variability being observed---it is even possible for miniscule changes in the configuration of the seafloor to affect the observations.

In practice, therefore, many techniques have been developed to attempt to estimate uncertainties using non-ideal data, at least as an approximation. For example, if two different systems observe the same thing simultaneously, any difference in the pair of observations has to be caused by the systems, and not by the thing itself. This is sometimes used to make paired observations where the statistics of the difference between two devices are examined. Effective techniques can vary according to the

[2] Note that this is not the “true” depth, since the physical measurement also has uncertainties. It might be, however, significantly lower uncertainty than the acoustic measurement and therefore considered sufficiently “true” for current purposes.

measurement; more details for suggested methods with respect to crowd-sourced bathymetry are presented in the following sections, and in Appendix X.

Data sufficiently rich to be interesting are often the result of a number of different observations, or the combination of core data from multiple different systems. For example, with bathymetric observations, the resulting depth is a function of at least a positioning system and an echo sounder, tied together by a common sense of time, and more sophisticated systems might include one or more sound speed sensors, and a motion sensor. Attempting to estimate the uncertainty of the entire system in one step can be problematic, and therefore it is common to estimate the component uncertainties associated with individual measurements, and then combine them together according to the overall equations that define how the data are combined. The techniques used to do so can be very involved, but essentially separate out into either formal mathematical models, or simulation-based models. The former techniques are very useful for making predictions about how the uncertainties are going to behave under different circumstances, for example to assess the sensitivity of the overall uncertainty on one of the contributory factors; the latter are mathematically much simpler, and therefore allow for more complex situations to be tackled, but are often more computationally intensive. Methods for combination of uncertainties are addressed primarily in Appendix X, but have a large influence on the use of the data from the database.

Irrespective of how the uncertainty is estimated or combined, the ultimate goal is to provide a succinct statement to potential users of the data as to the expected uncertainty. The most common way to do this is to provide a range of values in which the true value of the measurement is expected to lie, with a given probability. So, for example, a depth could be specified as being "between 12.3 and 14.2m, 95% of the time". Where the range is either known or assumed to be symmetric, it is often the case that the mean value and spread might be given, so that the depth might be specified as "13.25 ± 0.95m, 95% of the time". Whichever method is used, it is important to be clear as to the limits, and the probability, since both are required to unambiguously specify the expected spread of values being reported. For some cases, it might also be important to specify how the values are expected to be spread about the range given. In most cases, this is typically assumed to be Gaussian distributed, where the mean value is the most likely, and the values tend to cluster about this value strongly. In some cases, however, other distributions might be more usual, or it might be more prudent not to make an assumption about the distribution at all, and say that the values are equally likely to be at any position within the quoted range. A Gaussian distribution is normally assumed unless specified otherwise. This being the case, a 95% probability range is most commonly reported; if a symmetric range is specified, the limits for this are approximately two standard deviations[3] of the values being reported.

Although statistical descriptions of uncertainty are most common, they are not ubiquitous. For example, it might not be possible to directly specify the uncertainty of some of the measurements, or there might not be sufficient information (e.g., in the metadata) to provide a complete description of uncertainty. Under these circumstances, while statistical methods are still preferred, more qualitative

[3] And exactly 1.96 standard deviations.

methods might be required. For example, data might be qualified as "Poor", "Medium" or "Good" quality, based on a subjective assessment of the data collector. Or, data might be rated on a scale of 1-5 based on its correspondence with other observers in the area. Methods for assessing, manipulating, and interpreting these types of uncertainties are significantly less developed than for statistical uncertainties.

5.1.4 Consequences of Uncertainty

Effects of uncertainty on how we interpret data. Limitations to use (or alternative uses) depending on the uncertainty specified.

5.1.5 Uncertainty for Individual Observers

Focus primarily on the things that the individual observer can do to estimate and represent the uncertainty. Proper installation. Measurement of offsets. Calibration. Provision of metadata. Generation of data that can be used for calibration (i.e., for others to use, rather than for the individual observer to compute over). Guidance for what to tackle first (i.e., list of priorities for an individual observer). Checklist.

5.1.6 Uncertainty for Trusted Nodes

Focus on the things that the Trusted Node can and should do to include the uncertainty of the individual observers into data provided. Meta-analysis of multiple users to establish baseline uncertainties and remove dubious data. Calibration and measurement uncertainty estimation for individual users. Cross-calibration of multiple observers to provide confidence measures in the observers. Tracking of observer data over time to provide a measure of reputation. Feedback of results of meta-analysis to individual users in order to provide guidance for what they should be thinking about fixing/improving. Vetting of metadata from individual observers. Publication of discovery metadata for individual observers (or db?). Partnership with authoritative data holders to allow recent authoritative data to be used as a comparison point (offload of responsibility?) Guidance for what to tackle first (i.e., list of priorities for a Trusted Node). Checklist.

5.1.7 Uncertainty for Database Users

Focus on the things that database users have to be aware of when using the uncertainty estimates that they can procure from the database. Understanding the uncertainties that are provided in the database. Combination of uncertainties. Modeling uncertainty & methods to estimate. Proper statement of the overall uncertainty of estimates. Feedback mechanism to report on doubtful soundings/observers? Guidance on the relative magnitude of the various effects that could be observed (i.e., list of priorities for a database user). Checklist.

6 Data Contribution

Data collectors of CSB are strongly encouraged to provide their data to the IHO Data Center for Digital Bathymetry, so that the resulting global database of bathymetric data can be stewarded and shared with all IHO member states and the general public. Providing the data in consistent formats with sufficient metadata will allow users of this dataset to assess the completeness and overall usability of the data. This section details the process for contributing data and specifies the data formats required, and is therefore more prescriptive than the rest of this document.

6.1 Methods for Contributing Data to the DCDB

The DCDB currently supports contribution of CSB data via a network of 'trusted nodes', individuals or organizations that serve as liaisons between a like group of mariners and the data center. This model ensures clarity of requirements and data consistency, while minimizing the effort on individual mariners to participate in contributing data to the public cloud. In the future, the DCDB plans to expand its capability to support other models, including individual mariner contributions.

DCDB works with each trusted node on an individual basis to customize a streamlined process for contributing data. For example, file format, transmission protocol, minimum metadata fields, authentication method, data logging, data aggregation techniques are weighed to select the best option, with the expectation that these can be updated to respond to changes in technology and software.

6.1.1 Transmission Protocol

Data contributors have the option of hosting the data and making it available for retrieval by the DCDB using FTP or by pushing the data to the DCDB via an HTTP request.

6.1.2 Data Aggregation (file recording or breaking)

Data is collected using different approaches depending on the platform and instrumentation, which results in a variety of file sizes. Data transmission method is considered when agreeing on file size limitations.

6.1.3 Data Logging Rate

6.1.4 Authentication Method

The DCDB needs to ensure the integrity of incoming data streams, and so a unique key is assigned to each trusted node to authenticate the provider.

6.2 Data and Metadata Formats

All data contributions should conform to the data format and metadata standards described in the "Metadata" section of this document, unless otherwise approved by the Director, IHO DCDB.

Data are currently being accepted in two formats, GeoJSON and CSV with a JSON metadata header. Data should be delivered as point soundings with the required fields of latitude, longitude, depth, and time for each sounding.

6.2.1 GeoJSON

The GeoJSON Format Specification

GeoJSON validator

6.2.2. CSV

This XYZT format requires a separate metadata header in a JSON format. [add link to latest format]

6.3 Contributing Data

[add diagram of flow]

See Tables 3.1 and 3.2

Notes:

- Include a simple explanation of the preferred up-load protocols (trusted node, single observer), how data mining/viewing can be undertaken, what download protocols are in place.

[outline specific conventions, field formats, etc.]

- Address how to handle downsampled datasets from other crowdsourced bathy organizations? Is this extracted from an interpolated grid? What method was used? etc…
- Recommend (require?) BZip2 or gzip compression on JSON files
- Recommend (require?) HTTPS web API for submission
- Allow FTP(S) for submission
- Allow REST for HTTPS POST of XYZ file and JSON string for metadata
- Need to specify the ruleset for the JSON validation as text for publication with guidance document; give fields some numbering so that errors can be reported; keep validation checks as minimal as possible to avoid too many rejects.
- Provide information on how error conditions are going to be reported to the contributor.
- Need to define (for NCEI developers) the attributes on which the data can be filtered in the end-point map viewer that they are developing.
- Data formats downloaded from the DCDB
- The data format available to the public when downloaded from the DCDB website (https://www.ngdc.noaa.gov/iho/#csb) will nominally be provided in native GeoJSON, equivalent to what was submitted by the data contributors. *ASCII format, as well as ESRI Shapefiles, and Feature Classes within a file geodatabase might be potential options in the future.*

- There should be an explanation of the different data formats used by various sensors and systems. The suggested appropriate and preferred formats should be detailed and explained why these particular formats have been chosen.
- So as not to put some participants off, there should be details of what process will take place
- if data is submitted in a different format from those identified above.

7 The International Hydrographic Organization Data Centre for Digital Bathymetry (IHO DCDB)

7.1 IHO DCDB Overview

The IHO DCDB was established in 1988 to steward worldwide bathymetric data on behalf of the IHO Member States. The Centre provides a long term archive of, and access to, sonar datasets contributed by a wide range of mariners from across the globe. The DCDB is hosted by the US National Oceanic and Atmospheric Administration's (NOAA) National Center for Environmental Information (NCEI) in Boulder, Colorado. The Centre stewards global bathymetric data to international standards and provides it to the public freely and without restrictions. All data hosted by the DCDB is accessible online via interactive web map services. Spatial footprints of bathymetric data that is not publically available or hosted on other sites are provided with metadata for easy search and discovery.

7.2 Overview of CSB data flow through the IHO DCDB

[Add Diagram]

7.2.1 Data Discovery and Access

7.2.2 Accommodating CSB data

Notes:

- There should be a simple explanation for how data mining/viewing can be undertaken, what download protocols are in place.
- There should be a description of data flow (processing/validation/quality assessment of data) - data collector →trusted node→DCDB→user.[1] This will explain what happens to all collected data, regardless of source, and show to individuals the whole process and when they can expect to be able to see their data in the public domain and available from the database viewing.

8 Legal Issues

8.1 Introduction

Numerous issues concerning the legality of collected and publicly distributed crowdsourced bathymetry data have been raised throughout the maturation of this project. This chapter attempts to identify and provide insight (not necessarily advice) on each. Please note that the issues listed are not exhaustive.

8.1.1 Liability

Issues concerning liability might arise from some CSB contributors. What if a vessel runs aground and blames the data contributor for providing incorrect data? Is the data provider to be held liable? It is important to recognize that contributors are not compensated for providing their data to the IHO DCDB. Their contribution can be described as simply “for the public good” which might discourage any liability issues should they arise. Also, all CSB metadata records state: “These data are not intended to be used for navigation” providing additional reasoning that the data provider should not be held liable.

8.1.2 Logging in the Territorial Sea and Exclusive Economic Zone (EEZ)

May a vessel log data in another coastal State’s Exclusive Economic Zone, or territorial sea that will ultimately be made publically available without first requesting permission from that country?

Under international law, as reflected in the 1982 United Nations Convention on the Law of the Sea (DOALOS, 1997), the territorial sea falls under the sovereignty of the coastal State. Although every ship is guaranteed the right of innocent passage through the territorial sea, that right is subject to certain restrictions.

Article 19 (1) of UNCLOS states that “passage is innocent so long as it is not prejudicial to the peace, good order or security of the coastal State. Such passage shall take place in conformity with this Convention and with other rules of international law.” Article 19 (2) explains that passage of a foreign ship is considered prejudicial to the peace, good order or security of the coastal State if in the territorial sea it engages in a variety of activities. Of the twelve activities (a-l) listed, one might ask where data logging might fall under three activities:

> *(c) any act aimed at collecting information to the prejudice of the defence or security of the coastal State;*
> *(j) the carrying out of research or survey activities;*
> *(l) any other activity not having a direct bearing on passage.*

If so, then the ship collecting data within the territorial sea would not enjoy the protection of innocent passage and would be subject to the jurisdiction of the coastal nation in whose territorial sea it is operating.

Recognizing that much data gathering remains to be done, the IHO Member States at IHC-18 decided: *"…to progress whatever actions are required to improve the collection, quality and availability of hydrographic data worldwide, monitor and rectify possible deficiencies and shortcomings, cooperate with other international organizations and stakeholders as necessary, and to keep Member States informed on progress on this issue. … ".* That data need is the driver behind the IHO CSB Project. Chapter X of this document also describes that the purpose of data logging and subsequent contribution to a data repository (IHO DCDB), is to make this data readily accessible to the community for a range of uses, including nautical charting and global digital datasets. Therefore, coastal States should be assured that data collection is not to the prejudice of the defense or security of that State.

With respect to data logging in the EEZ, some Coastal states would likely consider that such an activity constitutes "marine scientific research" (MSR), in which case the activity falls under the jurisdiction of the coastal State. Not all countries agree with this view. For example, the United States' position is that this type of data collection in the EEZ (12-200 nm from the coastline) is not considered "marine scientific research" that requires the permission of the U.S. government. However, for States that disagree, restrictions might be required and data collection could fall under the control of the coastal State. Generally, coastal States require those conducting MSR to request permission six months prior to the activity.

The question of, "who is the flag state" may also arise if there is a dispute over innocent passage as the flag State would likely be advocating on behalf of its vessel.

8.1.3 Agreed Consent

The IHO's approval and endorsement of data logging and sharing might allow for consent from coastal States. That is, if the coastal States understand the activity and its benefits, and participate in the decision-making process, they might be more persuaded to allow the activity in their waters without prior permission on a case-by-case basis.

There is a potential precedent for this idea of international agreed consent. The Argo Project consists of a global array of more than 3,000 free-drifting floats (buoys) that continuously measures the temperature and salinity of the ocean, with all data being relayed and made publicly available within hours after collection (ref 2). When the Argo Project first began in 2000, many coastal States protested that these buoys were going into their EEZs and collecting data. In an attempt to reach a solution, "the Intergovernmental Oceanographic Commission (IOC) requested that its Advisory Body of Experts on the Law of the Sea (IOC/ABE-LOS) draft Guidelines. The IOC/ABE-LOS drafted non-binding Guidelines for the legal regulation of Argo Profiling Float Deployments on the High Seas in 2008 which require prior notification if the Argo floats enter the EEZ of IOC Member States and which give coastal States some control over the public distribution of sensitive information" (ref 3).

The experience and precedent of the acceptance of the Argo Program may want to be explored by the IHO in greater detail: Ref 4 and Ref 5.

8.1.4 Ownership of Data

Questions have been raised regarding whether or not there exists an "ownership of the data" therefore allowing the possibility of copyright establishment. It could be argued that since data logging is a direct observation of data, it might not be possible to establish copyright. One might compare the data to telephone numbers in a telephone book. The numbers themselves cannot be copyrighted. However, the phonebook itself could be. However, in this comparison, the "phonebook" would be the IHO DCDB, and CSB data set is not subject to copyright protection.

References:

1. Division for Ocean Affairs and the Law of the Sea (DOALOS), 1997, United Nations Convention on the Law of the Sea, New York: United Nations. 294 pp. (available at http://www.un.org/Depts/los/convention_agreements/texts/unclos/closindx.htm)
2. Argo Ref: http://www.argo.ucsd.edu/About_Argo.html
3. Mateos and Gorina-Ysern, "Climate Change and Guidelines for Argo Profiling Float Deployment on the High Seas."
4. J. A. Roach, "Defining Scientific Research: Marine Data Collection
5. Beckman and Davenport, "The EEZ Regime: Reflections after 30 Years" 2012

Appendix A – Echo Sounder Types and Technology

(TBD)

Appendix B – Data Contribution Format

Crowdsourced Bathymetry JSON
Format Version 1.0
Last Update: October 21st, 2015


```
{
  "type": "FeatureCollection",
  "crs": {
    "type": "name",
    "properties": {
         "name": "EPSG:4326"
    }
  },
  "properties": {
    "convention": "CSB 1.0",
    "platform": {
         "type": "Ship",
         "name": "White Rose of Drachs",
         "ImoNumber": "1008140",
         "platformStatus": "new",
         "draft": {
                   "value": 4.6[1] ,
         "uom": "m",
           “offsetApplied”: false
         },
         "sensors": [
                   {
                   "type": "Sounder",
                   "make": "Sperry Marine (L3 ELAC)",
                   "model": "ES155100-02",
                   "serialNumber": "136",

                   "longitudinalOffsetFromGPStoSonar[2] ": {
                             "value": 3.52,
                             "uom": "m"
                   },
                   "lateralOffsetFromGPStoSonar[3] ": {
                             "value": -0.76,
                             "uom": "m"
                   },
                   "velocity": {
                             "value": 1500,
                             "uom": "m/s"
```

```
                }
                },
                {
                "type": "GPS",
                "make": "Litton Marine Systems",
                "model": "LMX420",
        "offSetApplied": "false"
                }
        ]
    },
    "providerContact[1] Point": {
        "orgName": "Sea-ID",
        "hasEmail": "support@sea-id.org",
        "orgUrl": "http://www.sea-id.org"[4] ,
        "logger": "MarkIII",
        "loggerVersion": "1.0"
    },
    "processorContactPoint": {
        "hasEmail": "support@sea-id.org"
    },
    "ownerContactPoint": {
        "hasEmail": "support@sea-id.org"
    },
    "depthUnits": "meters",
    "timeUnits": "ISO 8601"
  },
  "features": [
        {
        "type": "Feature",
        "geometry": {
                "type": "Point",
                "coordinates": [
                40.9148,
                19.0052
                ]
        },
        "properties": {
                "depth": 15.8[5] ,
                "time": 2015-08-06T22:00:00Z[6]
        }
    }
  ]
}
```

Appendix C

(TBD)

Appendix D

(TBD)

Appendix E

(TBD)

List of Acronyms

List of Figures

List of Tables