HOW TO GUIDELINE

Supplement to S3/SC1.7-2026, AMERICAN NATIONAL STANDARD

Standard for Acoustic Metadata for Passive Acoustic Monitoring

Scenarios illustrating how standards can be used to record data associated with passive acoustic monitoring

Supplement to Standard for Acoustic Metadata for Passive Acoustic Monitoring

Working Group S3/SC1/WG 7, Acoustic metadata for passive acoustic monitoring, which assisted Accredited Standards Committee S3/SC1, Animal Bioacoustics, in the development of this standard, had the following membership at the time of submission:

Simone Baumann-Pickering – Scripps Institution of Oceanography, Univ. California, San Diego, CA, USA

Danielle Cholewiak, National Oceanic and Atmospheric Administration (NOAA) Northeast Fisheries Science Center, Woods Hole, MA, USA

Douglas Gillespie, St. Andrews University, St. Andrews, Scotland

Shane Guan, Bureau of Ocean Energy Management, Sterling, VA, USA

Jasper Kanes, Ocean Networks Canada

Katherine H. Kim, Greeneridge Sciences, Inc, Santa Barbara, CA, USA

Holger Klinck, Cornell University, Ithaca, NY, USA

Xavier Mouy, Woods Hole Oceanographic Institution (formerly of JASCO Applied Sciences, Victoria, BC, Canada), Woods Hole, MA, USA

Marie A. Roch (chair) – San Diego State University, San Diego, CA, USA

Ana Širović – Norwegian University of Science and Technology, Trondheim, Norway

Carrie Wall – University of Colorado Boulder and NOAA National Centers for Environmental Information, Boulder, CO

Table of Contents

1 INTRODUCTION

This annex provides a series of example scenarios that illustrate how the American National Standards Institute (ANSI) standards for acoustic metadata can be applied to capture information related to acoustic recordings. Information related to acoustic recordings consists of details about how, when and where the data were recorded, details about the calibration of these instruments, and descriptions of the signals contained within the recordings along with positional information about the producer of the signal.

We divide this motivational document into several types of information, each addressing specific needs:

- *Deployments* describes details about specific instruments that have been placed in the field to log acoustic data. The deployments section describes the information that needs to be retained for each instrument. Example information includes configuration information such as recording sample rate, equipment configuration, etc., as well as details about when and where the instrument was located.
- *Ensembles* permit the grouping of multiple instrument deployments into a coherent unit. The goal is to provide a reference for detections/localizations that combine multiple deployments for tasks such as beamformed detection or multi-instrument localization.
- *Calibrations* permit the preservation of acoustic calibration data.
- *Detections* permit retention of information about events that have been found in an acoustic data stream such as the presence of biotic, anthropogenic, or abiotic signals. We make a distinction between signals that were looked for systematically versus ones that were observed in an ad-hoc manner.
- *Localizations* provide for the preservation of mined location content. Location information can be in a variety of forms such as bearings, relative or absolute positions, and tracks.

Each of these data records has a list of subrecords and fields that contain the information relevant to representing information about these concepts. There are a number of record types that are used in multiple contexts. Recurring records are described in a common records and fields section (p. 31) at the end of this document rather than repeating the explanation each time. Sometimes, it is helpful to contextualize records or fields, and rather than writing Y is a subrecord or subfield of record X, the notation X/Y may be used.

This annex illustrates how the standard can be used as opposed to defining the standard itself. Please refer to the standards document for a complete list of records and fields. The standard specifies for each record or field whether they are mandatory, optional, or conditional, as well as the type of data that can be stored in fields.

For each major record type, it is recommended to read scenarios in order. Different scenarios frequently have common data requirements and in the interest of brevity subsequent scenarios primarily focus on the additional information that were not covered earlier.

2 DEPLOYMENT

We define a deployment as the placement of an instrument in the field. In general, we need to retain detailed information about how instruments were configured during a deployment, where they were deployed and recovered, and details about instrument components. We provide several scenarios that illustrate how fields of a deployment would be configured for different types of tasks. For widely distributed acoustic sensors (transducers, e.g. microphones, hydrophones), it can sometimes be difficult to know whether to consider them as one deployment or multiple ones. There is no hard and fixed rule for this, but generally transducers that are sampled using different oscillating clocks should be considered as separate deployments. One should not assume that separate deployments have synchronized sampling unless specifically noted in the Description subfield of the Methods record, and users of data from multiple deployments will generally need to account for separate clock drift if time-synchronization is required.

2.1 FIXED SINGLE-CHANNEL MOORING

A group deploys a single channel moored instrument such as a Rockhopper (Klinck *et al.*, 2020) or a high frequency acoustic recording package (HARP; Wiggins and Hildebrand, 2007) and wishes to capture metadata about the recording.

All deployments must be identified by a unique character string (Id). While this Id can be any valid set of characters, we recommend using meaningful information such as the project name, deployment site, and/or sequence number in a set of deployments. Remaining information falls into the following categories: where and why an instrument was deployed, identification about the recording instrument and associated sensors, and how data were sampled.

While the Id uniquely identifies the deployment, several fields provide information that can be useful in organizing deployments: Project, DeploymentNumber, and Site. Project is a text string whose contents are largely at the discretion of the user, but is typically used to denote a geographic region or monitoring project. Site is a descriptive name for the location. DeploymentNumber is an integer intended to identify an instrument deployment within a sequence of related deployments. It is used to denote the N^{th} deployment of an instrument relative to a reference of the user's choice such as the N^{th} deployment at a specific site or the N^{th} deployment of the project. Two optional convenience fields are provided for when alternate names exist for a Site or DeploymentNumber. The DeploymentAlias field can provide a textual alternative to the numeric DeploymentNumber. The SiteAliases record supports one or more Site fields that provide alternate names for a site.

An optional Description record within a deployment permits unstructured descriptive text that can answer many of the why/how questions about a deployment. Description can contain fields: Objective, Abstract, and Methods. The Description record is present in other types of records (e.g., Localizations) and is further described in the "List of common records and fields" section (p. 31).

The Instrument record is a general description of the recording equipment and consists of three fields with general information. The Type field is a text string that declares the type of instrument model or manufacturer (e.g., HARP, Rockhopper, Snap, AudioMoth, or SoundTrap). The InstrumentId field is a string that uniquely identifies the instrument and would typically contain a serial number. Field

GeometryType is optional and contains one of two values: rigid or cabled. This is mainly useful for multiple sensor configurations as it indicates whether the sensor geometry is likely to change over time which is important when the data are used for localization. In our example, let us assume that a hydrophone is attached to a line with a float above the data logger in which case we would state that it was cabled as it is likely to shift positions in the current.

Sensors is a record that contains a sequence of subrecords that describe the recording system's sensing devices. A subset of the ISO 19115 or OGC standards (International Standards Organization, 2003; Open Geospatial Consortium, 2023) appropriate for audio data and its consequent analysis is supported: Audio, Depth, and a generic Sensor. In a fixed single-channel mooring, the Audio record would be the only sensor record to be populated. Mandatory Audio fields are: Number and SensorId. Number is simply an integer that can be used to associate an audio sensor with a specific channel and sampling regimen that will be described below. SensorId is a string used to identify a specific sensor package, such as a serial number. A number of optional fields exist, for this scenario (and most others) TransducerId and PreampId are relevant. TransducerId can contain information needed to identify a specific microphone or hydrophone, and the PreampId identifies a preamplification unit. These are mainly useful when dealing with custom equipment where individual components need to be tracked and calibrated. When dealing with consumer-off-the-shelf hardware, one may not know much about the individual components of a unit. In these cases, one would still want to create an entry for the Audio transducer and only populate the Number and SensorId, making the Number and SensorId identical. Since the transducer is unlikely to be changed out in a commercial-off-the-shelf system, one can think of the sensor and unit as synonymous.

The SamplingDetails record contains a list of Channel records. Each Channel record contains information about one of the acoustic data channels. The record permits one to associate acoustic data with specific audio sensors and describes information that characterize the data stream such as when the data started, ended, the sample rate, duty cycle, etc.

A channel is associated with a specific audio sensor by specifying a ChannelNumber and a SensorNumber. The ChannelNumber, a non-negative integer, associates the channel recording properties with a specific stream of data within a recording. In this single-channel example, we might choose to use the index 0 for the channel. The data within that stream will have been acquired from the hydrophone Audio record that contains a Number that matches SensorNumber. In this single-hydrophone scenario, there should be exactly one Channel record and one Audio record.

Start and End fields contain UTC timestamps that indicate when sampling of the sensor began and ended for this channel. Note that sampling may begin before or after an instrument is deployed and is thus different than the deployment and recovery timestamps that will be discussed below.

Other fields within a Channel record specify sample rate and sample depth (bits/sample), gain, and duty cycle. For each of these (Sampling and the optional Gain and DutyCycle records), a sequence of Regimen records permit changes in parameters within a deployment. Each Regimen record consists of a UTC TimeStamp indicating when the specification becomes effective and information about the parameters. The Sampling record has Regimen records for sample rate and the number of bits used to quantize each pressure measurement (SampleRate_Hz and SampleBits). Gain regimens contain either calibrated (Gain_dB, reference intensity must be specified: usually 1 µPa for aquatic measurements, 20 µPa for terrestrial ones) or uncalibrated settings (Gain_rel) such as a number on a dial. Finally, the

DutyCycle record contains the number of contiguous seconds of recording (RecordingDuration_s) in an N second period (RecordingInterval_s). It is assumed that the recording period of the duty cycle is at the beginning of each RecordingInterval_s, but an optional Offset_s attribute permits this to be positioned arbitrarily within the recording interval. Recording intervals start at the specified TimeStamp. It is recommended to align recording intervals with day boundaries as it can make certain types of analysis simpler. In this scenario where we are recording continuously, we would produce a Sampling record with the specified SampleRate_Hz and SampleBits. For a continuous recording with no software-defined gain, we would not employ Gain or DutyCycle records.

Deployments have an optional QualityAssurance record that describes any steps taken to ensure data quality. Optional but highly recommended records within QualityAssurance are Description and ResponsibleParty, both of which are described in detail in the "List of common records and fields"section (p. 31). Description serves here to document the quality assurance process. ResponsibleParty is based on the ISO 19115 standard and contains numerous elements to identify an individual, their organization, and contact information. Notes on data quality appear in a mandatory list of one or more Quality records. Each Quality record must contain mandatory Start, End, and Category fields. Start and End are ISO UTC timestamps, and Category must be one of the following: unverified, good, compromised, and unusable. In this scenario with a single hydrophone, we assume that data have been manually verified and that there are no problems. The Description record would contain information on how this was checked, who checked it (ResponsibleParty), and a single Quality record that spans the range of the recording (Start/End) would contain Quality good. Had there been issues, additional optional fields are available for specifying the impacted frequency range, the channels affected, and any comments associated with the quality assurance process.

The Data record can contain a set of subrecords that provide additional information about how to access the acoustic data and geospatial tracks for moving instruments. In this scenario with a moored recorder, only information about how data should be accessed is relevant. Consequently, only one subrecord is populated: Audio. It contains a single mandatory field, URI, which is an abbreviation for uniform resource indicator, text that provides how to access an Internet resource (Berners-Lee *et al.*, 2005). Web addresses, or uniform resource locators (URLs) are a form of URI. When data are stored in a manner that is not network addressable, one may use descriptive text in this field (e.g., file cabinet 3) as it is simple for software to distinguish between descriptive text and a valid URI. Other optional elements are available for further information about the URL and to locate some derived data products and raw instrument data.

DeploymentDetails and RecoveryDetails permit recording information about the deployment and recovery, and both have identical elements. Mandatory elements include Position that records where the instrument was deployed and recovered using a standard coordinate system such as latitude and longitude in WGS84. The date and time of deployment and recovery are recorded in TimeStamp. TimeStampAudio contains the start or end of the recording effort which might differ from the deployment/recovery TimeStamp. ElevationInstrument_m, denotes the instrument elevation or depth below sea level (negative). Several optional fields related to this may also be used: Elevation_m, and DepthInstrument_m. Elevation_m specifies the elevation of the ground or seabed (negative) in meters which may be different from the elevation of the instrument itself (e.g., a recorder attached to a tree on a hill might have an elevation of 350 m but the ground is 345 m above sea level). DepthInstrument_m allows specifying aquatic instrument depth with positive numbers when the water surface is not at sea level, such as one would find in a mountain lake. Vessel denotes the vessel that deployed/recovered the instrument, and contact information may be specified using Person (surname, name, userID, affiliation, phoneNumber, email – a subset of the OpenGIS Sensor ML standard) or ResponsibleParty (ISO 19115) record.

2.2 FIXED ARRAY

A user wishes to deploy a moored instrument that has multiple hydrophones. For the most part, this is identical to the previous example with minor differences. Additional elements of type Sensors/Audio will be required to specify the multiple hydrophones and their configuration. We will consider a hydrophone array (Figure 1) that was deployed from the floating instrument platform (FLIP) (Gassmann *et al.*, 2013). FLIP is a moored platform that remains in a static position throughout a deployment via multiple mooring lines. As we describe the positions of the hydrophones, we will assume a coordinate reference system for FLIP that has its zero point at the center of the platform at sea level.

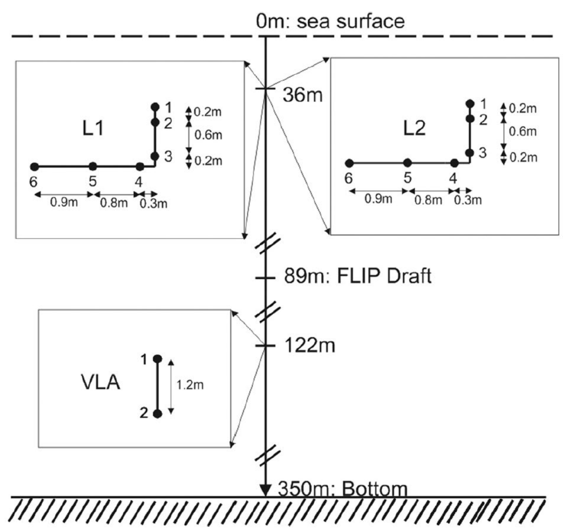

Figure 1- FLIP array deployment from Gassmann et al. (2013). The vertical linear array (VLA) and one of the two vertical and horizontal arrays (L1 or L2) were combined to form more complex hydrophone array.

For simplification, we will only discuss the L1 and VLA arrays and assume that the horizontal component of the L1 array runs perpendicular to the FLIP's coordinate reference system. The top of the L1 array was 36 m beneath the surface, 6 meters port and 14 m in front of the origin. L1's hydrophone 1 was thus at an offset of $L1_1$:$(-6,\ 14, -36)$. The VLA was directly below the FLIP and its top element had an offset of VLA_1:$(0, 0, -122)$. With these assumptions other positions can be calculated by adding the horizontal and vertical offsets to these transducers.

Each of the 8 hydrophones (6 in L1 and 2 in VLA) will have an Audio subrecord within Sensors. There are several differences with the single channel deployment. First, the choice of the Number field value within Sensors becomes more critical as each sensor must have a unique number that permits us to associate the sensor with a channel of the recording and a specific sampling plan. Secondly, while the Audio Geometry field is optional, recording this information becomes critical if localization analysis is to be done on the data from the deployment. The Geometry record consists of three fields, x_m, y_m, z_m, denoting the offset from the zero point of the instrument's coordinate reference system. Sensor 1 of array L1 would have a Geometry subrecord with fields: x_m=-6, y_m=14, z_m=-36. Sensors 2 through 6 would have offsets relative to $L1_1$ as per Figure 1. Sensor 1 of array VLA might be called sensor 7 and would have a Geometry subrecord (x_m=0, y_m=0, z_m=-122) with the last sensor, sensor 8, being 1.2 m deeper. Finally, one would need to populate the text field ReferencePoint indicating that the origin of the coordinate system was at the center of the FLIP at sea level.

Assuming recordings were done on all 8 channels, there would be 8 Channel records in the SamplingDetails record. These would associate the physical hydrophones recorded in the Audio records to specific channels of the recording as well as specify sampling plans for each channel through the various fields of the Channel record discussed in the previous example. In many cases, the sampling plans will be identical (e.g., the pressure waveform is sampled the same number of times per second on all channels), but channels are permitted to have independent sampling plans.

Note that the manner in which most audio data are stored does not permit mixed sample rates, quantization bits, etc., and thus it should not be very common to have independent plans. Passive acoustic monitoring data are, however, sometimes sampled in different manners, such as having transducers with distinct properties to sample low and high frequency domains. As a consequence, the standard permits this, and Channel records must be provided for every channel of recorded data, even if they are all the same.

To include the L2 array, another set of Audio sensor and Channel records would be added in the same manner.

2.3 Towed array

A survey design calls for towing a two-element linear hydrophone array equipped with a depth sensor behind a survey vessel. As with the previous examples, many of the elements will be populated in a similar manner. We will focus on describing the sensor geometry and sampling.

Here we assume the hydrophones are towed at a fixed distance relative to the location where GPS reading are taken throughout the deployment. Scenarios involving dynamic tow-cable length would require an extension to the schema A Sensors/Audio record will be populated for each of the two hydrophones. The Sensors/Audio/Number is an arbitrary number used to associate the hydrophone with a sampling plan defined in SamplingDetails/Channel while the Sensors/Audio/SensorId allows the identification of the specific hydrophone by recording unique information such as manufacturer and serial number. For simplicity, we will assign 1 to the SensorId of the hydrophone closest to the ship and 2 to the second one.

The Sensors/Audio/Geometry record allows the specification of relative distances from the trackline measurements. Assume that hydrophone 1 is N_1 meters behind the GPS coordinate and hydrophone 2 is N_2 meters behind N_1. As the hydrophones are towed behind the boat, these would be entered as $-N_1$ and $-(N_1+N_2)$ in the y_m portion of the coordinates with the x_m set to zero. As this deployment has a depth sensor, the z_m element may be omitted. Otherwise, it should be set to the target tow depth. Specification of time-varying depth information is discussed below in the discussion of the Data record.

The depth sensor would be handled similarly, with a Sensors/Depth record containing a minimum of Number, SensorId, and Geometry.

The SamplingDetails record is used to specify how each of the hydrophones are sampling. For each of the two hydrophones, a Channel record is created within SamplingDetails. Population of the Channel records is similar to the fixed array example, with details about sample rate, duty cycle, start and end times, etc., contained within the Channel records.

As we wish to capture positional information about the instrument, the Data record will contain an additional subrecord called Tracks. Tracks consists of several fields and subrecords detailing unit choice, individual Track records, optional specification of effort, and the potential to reference Internet sources for track data.

SpeedUnit specifies the units that are used to record speed if a speed is associated with a track measurement. It must be kn (knots), m/s, or km/s. Knot is not an international standards unit, but it is frequently used in aviation and nautical measurements where 1 kn is equivalent to 1.852 km/h.

A sequence of Track records represent a set of observations. Each Track record contains an optional TrackId, a number that identifies a group of point measurements. Some of the measurements require that units of measurement be specified. Field North indicates whether Bearing_DegN and CourseOverGround_DegN refer to magnetic north or true North, values north and true.

In addition to TrackId, Track contains a list of a Point records that specify measurements on the track. Each Point has a mandatory ISO 8601 TimeStamp. By itself, the TimeStamp is not particularly useful, but when coupled with optional elements that describe position, velocity, instrument orientation, and ground elevation or bottom depth, the deployment of mobile instruments can be tracked. The rationale for making all other measurements optional is that not all measurement devices will produce time synchronous measurements. In this scenario where we record a GPS location from the ship and a depth reading from the depth sensor, it is unlikely that these will be sampled at the same time. Thus, Point records are likely to have one or both of the following (additional fields not relevant to this example are defined in the standard):

- Position in a standard coordinate system
- Elevation_m –the number of meters beneath sea level

Additional fields in Tracks allow the specification of track effort including when the survey passes through regions of interest (e.g., a protected area) and details about where track data are stored if they are not stored in the Track records described above.

2.4 Cabled observatory

A cabled observatory has 5 acoustic sensors used to form a large aperture array. This is similar to the fixed array example and can be handled very similarly. The major difference is that there is no fixed recovery date as data collection is ongoing. If the observatory records continuously, one may wish to use the last available recording date as the recovery date and periodically update the recovery dates in the deployment record or record the planned obsolescence date. The network geometry would be specified in the Sensors/Audio/Geometry as above. The general assumption for a deployment is that all of the sensors are sampled using the same clock, meaning that one does not need to consider clock drift across channels. When sensors are widely spaced, this may not be true and it may be beneficial to treat the array as separate deployments.

Recording multiple deployments for the observatory has several advantages and is how the US Navy uses this standard to describe their acoustic ranges. When instruments are widely spaced, it is easier to simply use the latitude and longitude of each instrument and describe sensor geometry on each platform. In addition, replacement of failed instruments in ocean observatories is nontrivial and can require some time to replace. This is advantageous as the Deployment records let users know the precise time that each instrument was available. Using a single record for the whole observatory would require specifying failed recording periods in the QualityAssurance and these would have to be searched to select sets of deployments that were appropriate to use as an array during a time interval of interest.

2.5 Event-triggered recordings

Some recorders log acoustic data based on a sensor state. Examples include detector- and temperature-triggered recorders. Triggered recording events are typically characterized by an aperiodic nature. These are handled by populating the EventTrigger record, a subrecord of Deployment/SamplingDetails/Channel only used for event-triggered recording. As EventTrigger is a per channel structure, different channels need not use the same event trigger. EventTrigger contains two subrecords, Description and Algorithm, that are described more completely in the "List of common records and fields" section (p. 31). As an example, consider a deployment that only records when a machine-learning algorithm predicts that the likelihood of a call occurring exceeds a given threshold. The Description record could contain descriptions of the algorithm and why we are using it in the Objectives, Abstract, and Method fields. A more formal description would be required in Algorithm, which contains a minimum of the Software used along with any settable Parameters, but also includes many optional fields such as version numbers, additional software used, etc. Parameters is an unconstrained field that can have user-defined subfields. Candidates for inclusion in this scenario might be the threshold used, the number of seconds before and after the event detection that are recorded, etc.

Some event recorders do not record acoustic data and only note when an event has occurred. In these cases, the Data/Audio/URI field should be set to the value "No data", and this should also be indicated in the text within the Deployment/Description record.

Detections associated with such recordings are saved separately using the detection schema and follow the standard paradigm for detections.

3 Calibration

Calibrations of instruments or instrument components are kept separately from deployments. This permits the association of multiple calibration methods with the same deployment. Recording of calibrations is supported for end-to-end instrument calibrations based on calibration of recordings produced by a fully-assembled instrument or combining the calibrations of each stage of the instrument.

People producing calibrations for deployments should either use an end-to-end calibration, or provide both hydrophone and preamplifier calibrations as both contribute to the end-to-end calibration.

To avoid repeating "microphone or hydrophone" many times, the generic term transducer is used throughout this section. The calibration record contains the following fields:

- Id – An identifier for the hydrophone, preamplifier, or instrument. When possible, this should be the serial number. These should be matchable to elements in deployments to determine when and where a calibration should be used. Unlike Id elements for other types of information, this field need not be unique as an instrument may be calibrated multiple times. A calibration is uniquely identified by its Id and Timestamp.
- Timestamp – UTC date and time the calibration was performed.
- Process – An algorithm description consisting of multiple elements indicating how the calibration was performed. See the description of Method in the Detections example for details on the elements contained within.
- ResponsibleParty – Optional element with children that allow the identification of the individual and organization that performed the calibration (see page 30).
- QualityAssurance – Optional record containing a Quality field (unverified, verified, invalid), and optional fields for Comment (text) and an AlternateCalibration. AlternateCalibration should be specified when the calibration is known to have quality problems and provides the Id and calibration TimeStamp of a suitable alternative calibration that may be used as a proxy for this one.
- IntensityReference_uPa – Sound pressure in dB is always relative to a reference pressure. This element records the reference pressure in µPa. Standard values:

Medium	Reference intensity	Value
air (terrestrial)	1 V/20µPa	20
water (aquatic)	1 V/µPa	1

  The value here should be thought of as the reference pressure, P_0, in the standard dB equation from a measured pressure P: $10\log_{10}\left(P/_{P_0}\right)^2 = 20\log_{10}\left(P/_{P_0}\right)$ and is referred to as P_0 in other definitions.
- Sensitivity_dBV – Optional measurement of transducer sensitivity at a flat point in the frequency response curve measured in dB re. $\frac{V}{P_0\ \mu Pa}$. In theory, the measurement point should not be critical as the response is assumed to be flat (see SensitivityLow_Hz, SensitivityHigh_Hz for specifying the range limit).
- AnalogToDigitalSensitivity_dB – Optional sensitivity of the analog to digital converter in dB re. counts/volt. Counts are the integer values sampled by the analog to digital conversion process. Their range is defined by the number of bits, b, used to represent the number which in many systems is configurable, with common values being 16 and 24. Integer values for b bits range from $-(2^{b-1})$ to $2^{b-1}-1$. Note that some libraries for reading audio files may convert the integer values to real numbers, typically scaling the minimum/maximum range to [-1, 1). In such cases, this value cannot be applied without taking into account the conversion.
- Sensitivity_dBFS – Optional measurement that indicates dB relative to $P_0\ \mu Pa$ for the largest value, or full-scale, that a system is capable of sampling. In this case, sensitivity measurement is measured relative to peak output of the full-scale signal. As with AnalogToDigitalSensitivity_dB, full scale is relative to the integer counts, see above for discussion. Many audio reader libraries normalize audio samples to [-1, 1]. Multiplying normalized values by Sensitivity_dBFS will provide the calibrated dB levels.

- SensitivityLow_Hz, SensitivityHigh_Hz – Optional fields indicating the lower and upper limits (Hz) of the region over which the instrument exhibits a flat response (≤ ± 3 dB).
- FrequencyResponse – Optional measurements of the sensitivity in dB for various probe frequencies. Children are ResponseType, Hz, and dB.
    - ResponseType – Must be one of:
        - absolute – Response shows the absolute gain for each measured frequency taking into account all prior amplifications.
        - relative – Represents an offset relative to the flat portion of the curve represented by AnalogToDigitalSensitivity_dB or Sensitivity_dBFS.
    - Hz – list of frequencies at which the response was measured
    - dB – list of dB relative to the reference unit that correspond to the absolute or relative gain at the measurement frequencies contained in the corresponding entry of Hz.
- SystemNoiseFloor – Optional measurements of the minimum intensity signal that the instrument will measure at specific frequencies. Contains a list of frequencies (Hz) and minimal intensities in dB relative to the intensity reference unit P_0. These values are always absolute measurements and are not relative to instrument sensitivity.
- MetadataInfo – Optional element that contains information on who is responsible for maintaining and updating this record.

Calibrations are currently only defined for acoustic sensors (e.g., no depth sensor calibration).

3.1 Calibrating a hydrophone to a reference hydrophone

Given a well-calibrated reference hydrophone, we want to calibrate and an unknown test recorder in a salt-water tank. We loosely follow an example provided by Doug Gillespie and Mark Johnson, where we assume that a series of 500 μs pulses with frequencies between 5 and 40 kHz in 5 kHz steps are produced with a few seconds of silence after each pulse. The test sequence is repeated a second time. The short duration is designed to allow easy separation of the direct-path signal from echoes which are likely to follow quickly in a tank calibrated system. In this scenario, we assume that a reference transducer is used to generate signals and that the reference and unknown hydrophones are close together and .5 m from the sound source. The reference amplifier has an amplifier that provides a gain of 50 dB re 1 μPa. Both hydrophones/preamplifiers are connected to the same analog to digital converter and are recorded on separate channels.

The Id in the Calibration record will contain something to uniquely identify the hydrophone and preamplifier combination, and TimeStamp will contain the date and time at which the calibration was conducted.

The Process record describes the calibration process. The Method field contains text that indicates how the calibration was performed. In this case, we might write “peak-to-peak comparison method, Urick, R. J. (1983). *Principles of underwater sound* (McGraw Hill/Peninsula Publishing, Los Altos Hills, CA).” Additional fields are available to cite any software used in the calibration, these fields are identical to the Algorithm record used in detection and localization records (see examples in sections 4 and 5). The

Person record contains information about the person performing the calibration (name or other identifier and contact information) and the ResponsibleParty record contains similar information that describes who is responsible for the calibration. An optional separate record, MetadataInfo contains information about who is responsible for maintaining and updating this calibration record (Contact subrecord), an ISO 8601 timestamp indicating when the calibration was done (Date, e.g., YYYY-MM-DDTHH:MM:SSZ where Z indicates UTC time), and how often the record is planned to be updated. For calibration records, the value "as-needed" is usually the correct choice as calibrations are typically static unless an issue is identified.

The IntensityReference_uPa field indicates the reference pressure used in decibel computations. In-water calibrations should be relative to 1 µPa, and in-air calibrations relative to 20 µPa, so in this case we would use a value of 1.

To further populate the Calibration record, the calibration must be performed. For the peak-to-peak comparison method, we start by calculating the peak-to-peak gain for each frequency. We highpass filter the signals to easily identify the pulses and compute the peak-to-peak measurement for the direct path of each test signal. For any given segment of the recording containing a test signal, we compute the difference between the maximum and minimum sample value for both the test hydrophone and the reference hydrophone: $p2p\_test$ and $p2p\_ref$. The ratio of these two measurements can be converted to dB and is the relative sensitivity: $20\log_{10}\left({}^{p2p\_test}/_{p2p\_ref}\right)$. When multiple measurements are made at the same frequency, their ratio can be averaged before converting to dB to reduce measurement error. In this example, assume that for 5 kHz the relative sensitivity is -5.2 dB, meaning that at 5 kHz, any signal measured on the reference hydrophone will be -5.2 dB lower intensity on the test hydrophone. This is repeated for each frequency.

We are ready to compute the sensitivity of the test hydrophone using the sensitivity of the reference hydrophone, the known gain of the reference hydrophone, and the relative sensitivity between the two hydrophone/preamplifiers. As the reference hydrophone is calibrated, we examine its sensitivity at 5 kHz from the reference calibration. If a single number is given for the reference sensitivity, it is assumed that the response is flat across a specific range. Let us assume that reference sensitivity is -200 dB re 1 µPa across our calibration range. Given our reference hydrophone gain was 50 dB, the sensitivity of the test hydrophone at 5 kHz is:

$$\text{sensitivity: } (-200 + 50 - 5.2)\text{ dB re 1 mPa } = -155.2\text{ dB re 1 mPa}$$

We repeat this for each of the calibration frequencies. The record FrequencyResponse can be used to record the sensitivity at each frequency. It consists of two fields: Hz and dB. The Hz field is a list of the frequencies that were calibrated: 5000, 10000, 150000, …, 40000. The dB field contains a list of sensitivity values corresponding to each of the test frequencies. In this example, we are computing the absolute sensitivity at each frequency and a field within the FrequencyResponse would note that these measurements are absolute as opposed to being relative to the sensitivity of the assumed flat frequency response.

3.2 RECORDING A MANUFACTURER CALIBRATION

This example relies on a manufacturer's published calibration. The SoundTrap ST500 (OceanInstruments[NZ], Auckland, NZ) provides an end-to-end calibration for each instrument based that can be retrieved from their web site by serial number. As an example, a typical ST500 has a recorder sensitivity of -1.8 dB re 1µPa/V and a hydrophone sensitivity of 177.0 dB re 1µPa Adding the recorder and hydrophone calibrations together, there is a 175.2 dB re 1µPa sensitivity.

This is a full-scale calibration, meaning that for sample values that are normalized between -1 and 1, scaling by the calibration value provides the received dB re 1µPa.

To record this calibration, we would use the serial number of the instrument in the Id field. The Timestamp field would be populated with the timestamp that the manufacturer performed the calibration. OceanInstruments provides the calibration date through their web portal. For manufacturers that do not provide a calibration date use something appropriate such as the release date of the instrument or purchase date. The Process record can be fairly simple and need only populate the fields that we can be certain about. In this case, we might only populate that the Method was a piston phone calibration, and note in the Parameters that the only frequency tested was 250 Hz (as documented in the SoundTrap ST500 manual).

In the 2018 version of OceanInstruments' user manual, they do not describe the range over which the response is flat. Consequently, we would not be able to provide the optional SensitivityLow_Hz and SensitivityHigh_Hz, values showing the frequency range over which the response should not vary by more than ±3 dB. Wiggins and Morris (2019) report personal communication with OceanInstruments and indicate that response should be flat between 20 to 60,000 Hz and provide detailed frequency response calibrations for four ST500 units. They also report information on beam patterns which are not covered in the current version of the standard.

3.3 MISSING/CORRUPTED CALIBRATIONS SCENARIO

Calibrations are sometimes unavailable (e.g. calibration has been lost and equipment is no longer available to recalibrate or may have changed) or we learn that there was an error in calibration. The optional QualityAssurance field permits the specification of quality information. In the case of a bad calibration, it will be marked as invalid. The person examining the record could look at the QualityAssurance/Comment to find any details about the problem and QualityAssurance/AlternateCalibration for the Id of a different calibration that might serve as a viable alternative.

4 DETECTIONS

Detections refer to the identification of acoustic events or measurements of interest within a set of acoustic data. These may be animal vocalizations, anthropogenic sources, or abiotic events.

4.1 General requirements

Regardless of what we are trying to detect, there are a number of mandatory fields that are required to make a set of detections useful. We need to identify information about the recording or set of recordings being analyzed (the deployment or ensemble document), what methods were used to create detections, what they were looking for, and what was found. These are elaborated on through a series of examples. As always, refer to the standards document for an authoritative list of required and optional fields.

4.2 Specific species scenario

A researcher is interested in determining the distribution of Pacific white-sided dolphin (*Lagenorhynchus obliquidens*) echolocation click types. Soldevilla *et al.* (2010) identified two distinct types of clicks distinguished by distinct notch and peak spectral patterns. We assume that a detector/classifier has been trained that is capable of detecting clicks produced by this species, determining which of the two click types is being produced (if any), and assigning a confidence for the hypothesis with values in the interval [0, 1]. For each deployment that is analyzed, a separate set of detections will be produced.

Such a detector will at a minimum produce the following data for each echolocation click that is assumed to have been produced by a Pacific white-sided dolphin:

- Start time of click – either as a timestamp or an offset into the file. The standard requires UTC timestamps (date and time). Detectors that provide offsets into the file will need to be modified to produce these or have postprocessing software to translate sample indices or time offsets to timestamps.
- Click type A or B if appropriate.
- Confidence score.

Additionally, one might expect a call end time to be produced. While in this scenario, we are discussing the recording of every echolocation click, toothed whales can produce very large numbers of these and analysis of large data sets may be more manageable if one uses one of the standard's data summarization techniques such as reporting the number of echolocation clicks over a fixed period of time or per acoustic encounter (group of detections separated from other detections by a user-settable amount of time).

The schema for detections has additional fields that will need to be populated.

4.2.1 Mandatory fields

- DataSource – This allows the detections to be associated with an instrument deployment, permitting us to know location (including tracklines for mobile recorders), sampling rate, duty cycle, etc. Two options are available: one can either specify an ensemble of instruments or specify a specific deployment. Either one references the unique Id element in an ensemble or deployment record. Ensemble identifiers (EnsembleId) group deployments together into a virtual instrument. Most detection metadata do not need to use an ensemble unless it is for convenience or if multiple hydrophones from separate deployments are beamformed to enhance signal to noise ratio. Deployments are referenced using a DeploymentId field which contains a value which will match the Id of the associated deployment.

- Algorithm – The algorithm record contains data that permits one to reproduce the detection process. It consists of several subrecords, some of which are optional. The mandatory elements are:
    - Method – a textual description of the method which might cite a published algorithm or contain a summary description.
    - Software – the name of the software detector. For algorithms that are integrated into a general purpose passive acoustic monitoring package such as Ishmael (Mellinger, 2001) or PAMGAURD (Gillespie *et al.*, 2008), this is the name of the specific algorithm that may have been packaged as a library or "plug-in."
    - For reproducibility, it is important to know the values assigned to any user-adjustable settings in the algorithm. These are stored in the Parameters element and consist of text describing the parameter name and their value. Any number of parameters and values may be assigned, and they may be nested. For example, spectrogram parameters may be specified by the arbitrary parameter name "Framing" with child elements specifying the frame length and advance parameters. We encourage the use of structured text such as extended markup language (XML) or Javascript object notation (JSON) as data in elements that require open-ended data as standard libraries exist to generate and parse these specifications. As an example, a detector might use the following set of parameters expressed as XML:

      ```
      <Framing>
        <Advance_ms>2</Advance_ms>
        <Length_ms>8</Length_ms>
      </Framing>
      <Threshold_dB>10</Threshold_dB>
      …
      ```

      While it is recommended that one use a descriptive name for each setting, if the algorithm uses command line style parameters, e.g. -f 8 -a 2 -t 5, and it is desirable to store a string that can be extracted verbatim, we recommend that the structured text indicate that this is a parameter string, e.g.
      <ParameterString>-f 8 -a 2 -t 5</ParameterString>
      Using a structured representation make the search for algorithms with specific parameters easier and is recommended when feasible.

  The optional elements of Algorithm are:
    - Version permits a text specification the software release. When algorithms have version numbers, use of this field should be considered mandatory although it cannot be enforced automatically. In the absence of a version number other identifying descriptors, one may use a version control commit identifier (e.g., a specific commit checksum in a git repository).
    - SupportSoftware is used when the algorithm has dependencies. As an example, a detector that is written as a plugin package for PAMGUARD might rely on a specific version of PAMGUARD. SupportSoftware can be used multiple times to denote dependencies on other software packages. It supports three child elements: Software

(mandatory), Version, and Parameters that have similar semantics to the Software element described above.

- UserId – An identifier for the user that prepared the data that were submitted.
- Effort – A specification of the species and calls for which systematic search was conducted.
    - Child fields Start and End contain vectors of timestamps that specify the begin and end dates and times of the data that were analyzed. These vectors must be in ascending order and each End must be prior to the following Start. The earliest Start and latest End must fall within the recording effort specified in the deployment or ensemble. Users may elect to only analyze a subset of a deployment or ensemble. There is no need to specify multiple starts and ends due to duty cycled data as that information is derivable from the deployment information.
    - The Kind record may be repeated and indicates what the algorithm was looking for. It will contain field Taxon, a number based on the integrated taxonomic information system (ITIS Organization, 2014) representing Pacific white-sided dolphins, and element Call, which will be Clicks. As we are detecting click subtypes, two Kind records will be produced that are identical except for the Subtype field, one with Subtype A and the other with Subtype B. Finally, the mandatory element Granularity will be populated with “call” to indicate that the algorithm reports detections of individual calls as opposed to presence-absence over a defined duration (e.g. every 15 m) or groups of calls referred to as encounters.
- OnEffort – This is where detections that fall within the scope of the effort are reported. It consists of zero or more Detection records. In this scenario, the Start, End, Taxon and Call fields will be populated. In addition, a Parameters record will be will contain the Confidence field and the Subtype field when the system identifies an A or B click.

4.2.2 Optional fields

- Description – This is a high-level textual description of the detection process. Example for this scenario:
    - Objectives – Detect Pacific white-sided dolphin echolocation clicks and label ones that show evidence of A or B types peak and notch spectral patterns (Soldevilla et al. 2010).
    - Abstract – This is part of an effort to examine the tempospatial distribution of Pacific white-sided dolphin A and B echolocation clicks in the Southern California Bight.
    - Method – A support vector machine classifier was trained with data from three locations within the Southern California Bight distinct from this deployment. Clicks are detected with a Teager energy detector and classified via the support vector machine.
- QualityAssurance – This record has Description and ResponsibleParty. Description has an identical structure to the optional Description element described above. Its usage differs in that it describes the quality assurance process as opposed to the detection process. ResponsibleParty is described in the list of common elements (p. 30).
- OffEffort – This for reporting detections that were not part of the systematic effort and is usually only used with human analysts for reporting interesting events such as an unexpected call.

4.3 ANALYST SCENARIO

An analyst is using annotation software (e.g. Ishmael, PAMGUARD, Raven, or Triton logger) to look for minke whales (*Balaenoptera acutorostrata*) in Eastern Pacific waters off the Washington coast. She wishes to study boing call characteristics and would like to record the inter-pulse interval between each pulse in the call. There is no standard method for doing this. In addition, she comes across 3 calls from North Pacific right whales (*Eubalaena japonica*). While she has not looked for these systematically, their rarity makes preserving this discovery relevant.

Many of the fields that need to be stored are very similar to the previous scenario. DataSource, Algorithm, UserId, and Effort must all be populated appropriately. When humans are responsible for the categorization decision (as opposed to validating a machine decision), we recommend populating the optional Algorithm/Method field with the term “Analyst,” and the software field with details of the program that the analyst used to perform the annotation, capturing details of how the software was configured (e.g., spectrogram settings) in the Parameters element. In the Effort element, the user would not list the North Pacific Right whale as there was no consistent effort to find them.

The OnEffort element would have a list of Detection records with the Minke boing calls that the analyst detected. Start, End, Taxon, and Call fields would be populated. The UserDefined field in the Detection record’s Parameters record is designed to record data that have not been standardized. Like the algorithm parameters discussed in scenario 4.2, we recommend using structured text such as XML or JSON for such fields. As an example, one could place the following XML text in the USerDefined field:

<InterPulseInterval_ms> 14 14 14 14 14 … 14 11.3 … </InterPulseInterval_ms>

Finally, the North Pacific right whale detections would not be stored in the OnEffort section as they were not performed in a systematic manner. Those detections would be stored in a similar manner, but in Detection records associated within the OffEffort record.

4.4 SOUNDSCAPE SCENARIO

An analyst wishes to report a series of soundscape measurements across a series of 1/3 octave filters in a marine sanctuary. Measurements are estimated for every hour over a fixed time period associated with a specific deployment.

As with other scenarios, DataSource would need to be populated with an identifier that denotes the instrument deployment with which these measurements are associated. In the Algorithm section, we would populate Method to indicate how the measurements were computed, and Software to specify the software that computed the measurements. Parameters are algorithm specific, but in this case we will assume that the software designed proposed an OctaveStep parameter. As in other scenarios, we can use a structured XML/JSON text in Algorithm/Parameters, for example: <OctaveStep> 1/3 </OctaveStep> (XML) or { "OctaveStep": .333 } (JSON).

To denote effort for a soundscape, we would specify that any dB measurements in this set of detections are with respect to 1 µPa (Effort/IntensityReference_uPa = 1). Effort/Kind is used to specify our detection effort. In this case, we would set Kind/Taxon to -10 which denotes other phenomena (we use negative values to specify abiotic sources that are not anthropogenic). We would set Kind/Call to Soundscape, and populate Kind/Parameters/FrequencyMeasurements_Hz to a list of filter bank center frequencies. As these measurements are made hourly, we would set Effort/Granularity to binned with a BinSize_m of 60.

For each hour, we would generate a Detection with the Start/End showing the bin time span. The Detection would contain a Call (Soundscape), and Parameters/FrequencyMeasurement_dB that would be a list of measurements at each of the center frequencies.

4.5 Recording details of tonal calls

A number of software packages are available for following the frequency contours in tonal calls. These detections follow the Specific Species Scenario above (p. 17) and take advantage of the Detection/Parameters section of the Detections schema. A Detection/Parameters/Tonal field has three children that permit the specification of a moan, whistle, or other tonal call:

- Offset_s - A list of offset times in seconds relative to Detection/Start, representing the time at which measurements were made.
- Hz – The frequencies at which the measurements were made.
- dB – An optional intensity measurement in dB for each Offset_s/Hz pair. If the optional Effort/dBReferenceIntensity_uPa is present, these dB measurements are absolute, otherwise, they are assumed to be relative.

5 Localizations

While there are arguments for treating detections and localizations together, doing so makes managing different localization algorithms based on the same detections difficult. Consequently, the standard manages localizations separately from detections with mechanisms that permit the association of a localization event with one or more detections. This permits one to create multiple localizations records with different characteristics, each having its own identifier.

Some organizations will choose not to record the detections from which localizations were derived. While discouraged, this is permitted. A species identifier field, Taxon, can be used to note the species associated with the Localization in these cases.

5.1 General requirements

Each set of localizations have a text identifier (Id) that is unique with respect to other localization sets. As with detections, the fields Description, DataSource, Algorithm, QualityAssurance, UserId, BespokeData and MetadataInfo are used and serve the same purpose of identifying the data source based on the instrument deployment or deployments and how these data were processed and maintained. See the discussion of these elements in the scenarios described in the Deployment and Detection sections of this document.

An Effort section describes what types of location information are stored within this group of localizations. Like the detection effort, paired vectors of date and time timestamps delineate the start and the end of the localization effort. This is followed by a specification of a coordinate system.

CoordinateReferenceSystem is a set of data that defines the coordinate reference system used to denote location information within a set of localizations. A coordinate reference system is defined by a principal Subtype that indicates the type of coordinate reference system and a Name that specifies a specific reference system. We support a subset of the ISO 19111:2019 standard with one extension to support density estimation (measurements of perpendicular range to a track line). Translation of information to the ISO standard is straightforward. Valid values and their interpretation are summarized in Table I and more fully described below.

The Subtype field defines the type of coordinate reference system. Geographic indicates coordinates based on ellipsoidal projections, taking into account approximations of the Earth's curvature. Derived is any coordinate system that is derived from another one. Finally, Engineering is used for measurements in a Euclidean space.

The Name field specifies a specific coordinate reference system (CRS) within a Subtype. We support one CRS for Geographic (world geodetic system 84, WGS84), and one for Derived (universal transverse Mercator, UTM). WGS84 is the system used to provide longitude and latitude measurements by the global positioning system and should be familiar to most biologists/bioacousticians. UTM is a system that divides the Earth into local zones anchored to a WGS84 position with the assumption that within the zone a Euclidean frame provides a good approximation of position. Positions are specified by easting and northing measurements in meter relative to the anchor. The anchor is translated such that the measurements are always positive, and frequently called false eastings and northings due to this translation.

A number of coordinate systems are supported by Subtype Engineering. Cartesian measures offsets along three axes in meters. Polar measures a counterclockwise angle relative to an axis in a plane and potentially a distance. Spherical is similar to Polar, but adds a second angle that measures azimuth.

Table I – Coordinate reference systems and valid localization records

Subtype	Name	LocalizationType	Valid Localization Records
Geographic	WGS84	Point	WGS84 $(Lat, Long, elevation)$
		Track	Track of WGS84 vectors
Derived	UTM	Point	UTM $(E, N, elevation)$
		Track	Track of UTM vectors
Engineering	Cartesian	Point	Cartesian (x, y, z)
		Track	Track of Cartesian vectors
	Polar	Bearing	Bearing $(\angle_1)$
		Point	Angular $(\angle_1, d)$
		Track	Track of Bearing or Angular vectors
	Spherical	Bearing	Bearing $(\angle_1, \angle_2)$
		Point	Angular $(\angle_1, \angle_2, d)$
		Track	Track of Bearing or Angular vectors
	Cylindrical	Point	Cylindrical $(\angle, d_\angle, d_\perp)$
		Track	Track of Cylindrical vectors
Any	Any	Range	Range
		PerpendicularRange	PerpendicularRange

Finally, Cylindrical consists of a polar angle with a pair of distances. The first distance, $d_{\angle}$, measures the distance in the plane along the angle, and the second distance, $d_{\perp}$, measures an offset to the plane.

Coordinate systems must have some type of reference frame, or datum, that defines the origin of the coordinate system. Geographic subtypes have an implicit origin, but others define their origin in Subtype-dependent values in record ReferenceFrame.

For Derived UTM fields, ReferenceFrame contains an Anchor that consists of the value UTMZone and field UTMZone. UTMZone contains text with a zone number in the range of 1 to 60 which positions the longitude of a transverse cylinder positioned along the Earth followed by an N or S indicating whether UTM positions are relative to points in the northern or southern hemisphere. UTM is not usually used for high latitude locations where distortion is high.

For Engineering coordinate subtypes, the ReferenceFrame contains an Anchor which has one of the two following values: WGS84 or instrument. WGS84 indicates that all measurements are relative to a fixed point whose latitude and longitude are recorded in fields Latitude and Longitude within the ReferenceFrame. When instrument is used, text may be provided in the Datum field that provides a description of the zero-point relative to the instrument. Examples of Datum values: Location of GPS sensor, instrument center of mass. Additional text can and should be used to specify any other information needed to understand the measurements that will be recorded. Mobile instruments frequently have location tracks stored with the instrument, and of appropriate resolution these tracks may be sufficient to translate relative offsets to longitude and latitude measurements at a later date if so desired.

Examples:

- Engineering / Spherical – Localizations will be pairs of angles relative to axes defined in the reference frame and an optional distance.
  ReferenceFrame: Anchor=instrument, Datum=center of mass; angle 1 measured between center of mass and vector to nearest starboard side, angle 2 measured between center of mass and elevation relative to the transverse plane that contains the vector used to define angle 1.
- Derived / UTM – Localizations as false eastings, false northings, and possibly elevations relative to the specified UTM zone.
  ReferenceFrame: Anchor=UTMZone, UTMZone=31N
- Engineering / Cartesian – Localizations in meters east and north of the specified WGS 84 reference point.
  ReferenceFrame: Anchor=WGS84, Longitude=34.029, Latitude=32.06.
- Engineering / Cartesian – Localizations in meters along specified axes relative to the specified instrument frame of reference.
  ReferenceFrame: Anchor=instrument, Datum=location of inertial motion unit (IMU), x axis is along the vector from the IMU to the right/starboard side of the instrument, y axis is along the vector from the IMU to the bow/front of the instrument, and z axis is the cross product of these two vectors.

The next portion of the Effort record indicates what type of information is being saved regardless of the reference system. The LocalizationType field must contain one of the following values:

- Bearing – Localizations reflect a direction towards a sound source, but not the distance. Can be used with Name ∈ {Polar, Spherical}.
- Range – Straight line distance from instrument reference point to the sound source with no orientation information. May be used with any Subtype and Range.
- PerpendicularRange– Perpendicular distance to a track line associated with the instrument's motion vector which is assumed to be a straight line. PerpendicularRange is useful for distance sampling methods although this can usually be derived rather than stored. May be used with any Subtype and Range.
- Point – Each localization is expected to be a position. Used with any coordinate system that represents a position in two or three-dimensional space, and points may be represented in non-Cartesian coordinate systems such as a Polar or Spherical coordinate system.
- Track – Collection of timestamped positions that track a sound source's location over time.

In most cases, localized signals rely on detections from multiple channels. Differences in propagation paths result in different arrival times, resulting in an ambiguity in reporting the time associated with a localization. The standard defines field TimeReference within the Effort record to specify how an algorithm reports timestamps. TimeReference must contain one of the following values:

- absolute – Delays and position were used to estimate the time at which the sound was produced.
- channel – All times are presented relative to arrival on specified channels. Individual localizations will contain information that associates the localization with a timestamp relative to a specific channel. Details for each localization record (defined below) involve fields TimeReferenceChannel and potentially TimeReferenceEnsembleUnit.
- relative – Times may be relative to any channel of opportunity within the recording. For most applications, this is the recommended value.

The Dimension field contains a number that indicates the expected number of dimensions for each localization measurement, recognizing that there may be times where it is only possible to obtain a lower order estimate (e.g., a signal not detected on enough transducers to estimate elevation). Example dimension values:

- Geographic/WGS84 – Use 2 if estimates only contain latitude and longitude, 3 if the localization algorithm is attempting to determine elevation.
- Engineering/Polar – Use 1 if only estimating angle between the instrument reference vector and the sound source, 2 for angle and distance.

Use of this field makes it easy to distinguish between certain types of effort, such as effort to localize within a plane versus three dimensions.

The final Effort field, SpeedUnit, is only needed if instrument telemetry that measures speed is recorded. It denotes the units in which speeds were measured. Valid values: kn (knots), km/h, m/s.

The localizations themselves consist of a list of Localization records nested in a Localizations field. Each Localization record contains a series of fields. The first is an optional Event field that contains text to identify the localization. It must be unique with respect to all localization events associated with a specific Id.

A mandatory TimeStamp (ISO 8601 date time) represents the time at which the location information was estimated. Its interpretation is dependent on the TimeReference field in Effort. This is followed by an optional Taxon field (ITIS taxonomic serial number) that only need be used when the localizations are not associated with detections. The actual location information is recorded structures related to the type of measurement being made and consist of two records: Coordinate and CoordinateError. The Coordinate contains the direction or position and the optional CoordinateError provides the standard error of the measurement. The structures and the names of the fields in the Coordinate and CoordinaeError subrecords are:

- WGS84 – Longitude, Latitude, Elevation_m
- UTM – Easting, Northing, Elevation_m
- Cartesian – X_m, Y_m, Z_m
- Bearing – Angle1, Angle2
- Angular – Angle1, Angle2, Distance_m
- Cylindrical – Angle1, PlanarDistance_m, Elevation_m
- Range – Distance_m
- PerpendicularRange – Distance_m

An additional record Track, is used to denote location information related to a moving sound source. It contains positional or directional records similar to the ones above, but each of the measurements is a vector denoting a sequence of measurements. Timestamps contains a vector of dates and times that correspond to each location measurement. Finally, WGS84, UTM, and Cartesian coordinate systems, an optional Bounds record provides a bounding box around the detections to enable more efficient search for tracks.

The last field of each Localization record is an optional QualityAssurance record which accepts one of the following values:

- unverified – Validation has not been conducted. If the QualityAssurance field is not present, one should assume that the localization has not been verified.
- valid – Localization has been validated according to the specified QualityAssurance process defined for this set of localizations.
- invalid – Localization has been found to be erroneous according to the specified QualityAssurance process defined for this set of localizations.

5.2 Towed array: bearing from ship to animals scenario

It is common to tow multiple hydrophones in a line array behind a ship. In such cases, time-delay-of-arrival between signals at the different hydrophones can be used to estimate a left-right ambiguous bearing measured in degrees relative to the array. In this scenario, we will assume that the goal is to record the angle between the array orientation and the group of animals.

Directional information is stored in a localization document. As with all documents in this standard, an identifier field (Id) provides a sequence of characters that uniquely identifies the set of localizations. Any sequence of printable characters (alphanumeric, punctuation, space, hyphen, etc.), is valid. Common methods to specify an Id include some type of summary information about the instrument, deployment location or cruise, algorithm name, analyst initials, etc. The exact format is not critical as long as it is unique amongst the collection of localizations. Ids that are interpretable by humans are advantageous, but not required.

This is followed by the following fields which have been discussed in the context of detections: Description, DataSource, Algorithm, QualityAssurance, ResponsibleParty, and UserId. These are used in the same manner as with detections and provide textual descriptions of what one is doing, what data are being analyzed, how the data have been processed, what type of quality control process was used, and who executed or is responsible for the analysis.

5.2.1 Effort specification

Specifying the extent of our localization effort is important and stored within the Effort record. Effort contains the same Start and End fields as in the detections documents to indicate what portion of the data associated with the data source have been analyzed. Effort also contains information on what type of localizations were created, what type of coordinate system was used and how date and time were reported.

In this example, we are using a two-element hydrophone array to estimate bearings between the array and a group of vocalizing animals (Figure 2). This will be a single counter-clockwise angle between the current array orientation and the animals. In a two-element array, the angle will be ambiguous (α or β) and we always report the smaller of the two. The standard does not maintain multiple hypotheses, and if there is ambiguity it should be noted in the Description and/or Algorithm.

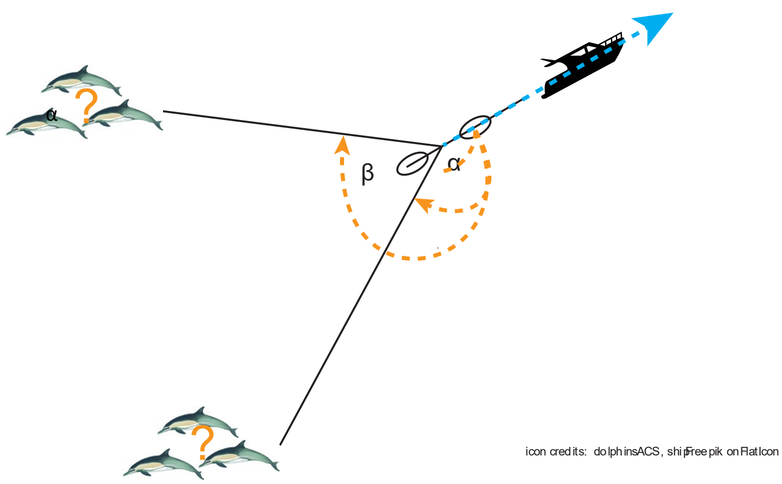

Figure 2 – Bearing angle from towed hydrophone pairs. Dashed blue line shows the vector relative to which bearings are measured. Ambiguous angles shown in orange, only the smaller angle will be recorded.

The CoordinateReferenceSystem field defines how our data will be reported. Since we wish to store polar coordinates, the CoordinateReferenceSystem will have a Subtype "Engineering" and Name "Polar".

As these polar coordinates are always relative to our acoustic array position, we note this in the ReferenceFrame record. Within the ReferenceFrame, we define an Anchor and a Datum. We set Anchor="instrument," indicating that all measurements are relative to the recording instrument or ship. The Datum would contain text indicating how the angle was measured, e.g., "Angle 1 measures the angle between the object and the array line that is assumed to be oriented in a line from the last hydrophone to the port stern of the ship."

The LocalizationType field will be set to "Bearing," indicating that we are estimating an angular direction towards a source. The Bearing label is used for any type of angular measurement when no distance in the specified direction is calculated. If a range had been computed, the value "Point" would be used as a point in $\mathfrak{R}^2$ space would have been specified. Point indicates that a point in space has been identified, regardless of the coordinate reference system used to identify it.

As we are only recording a single angle, we set the Dimension field to 1.

In this scenario, the array elements are close enough to another that we do not care which arrival time we record, so the Effort record's optional TimeReference field will be set to the value "relative." As animals are unlikely to move significantly between the times of arrival on the arrays, our choice of time of arrival is not important and we can pick one of the times or use a function of the arrival times. If the arrival time were critical, we would use value "channel," and report the channel whose timestamp we used with each localization.

5.2.2 Localizations

With the effort specified, we now consider the bearing angles that have been produced. These are stored in the Localizations field as a list of Localization records.

Each Localization will have several fields. An optional Event field lets us identify this localization uniquely within the document. A TimeStamp field provides a record of when we detected the sound at one of the hydrophones as Effort's TimeReference is set to relative.

In many cases, one can determine the species associated with the localization based on a referenced detection. If the user chooses not to link to a specific detection record, an optional Taxon field lets the user record a taxonomic identifier (taxonomic serial number, see detections for details). If the detections are referenced, they are contained in structure References which has the child field EventReferences. EventReferences contains a list of text value triplets: (type, id, event) where type specifies the reference type (localizations or detections), id the identifier associated with the localizations or detections, and event the Event label of a specific detection or localization.

Angle1 within a Bearing record's Coordinate stores the smaller of the two ambiguous angles between the reference axis and the vocalizing source. If the standard error was estimated, that is populated in Bearing/CoordinateError/Angle1.

If we wanted to post-process these data to make all angles relative to a single direction, we could use the orientation of the array over time to set up rotation matrices such that all polar angles were relative to a fixed angle. This would require knowing the orientation or heading of the array for each detection. Deployment records (discussed starting on page 10) have a mechanism for storing track information including information about bearing, speed, pitch, roll, and elevation. In practice, many people will downsample (decimate) track line data as they are often recorded more frequently than is needed for most purposes. As a consequence, the Localization record allows specifying these types of values with each localization so that high-precision data are preserved without extensive storage requirements. When this is required, an InstrumentTelemetry record can be used which contains optional information such as precise location in WGS84, pitch, roll, etc.

5.3 Localizations from towed array bearings

When multiple bearings are taken over time and if the animals have not moved significantly, it is possible to cross bearing angles to determine positions (Figure 3). There still remains a position ambiguity, but these can be resolved if the vessel turns during the transect or by use of selecting a hypothesis via a model selection method. The exact method used would be documented in the Algorithm section which has been previously described.

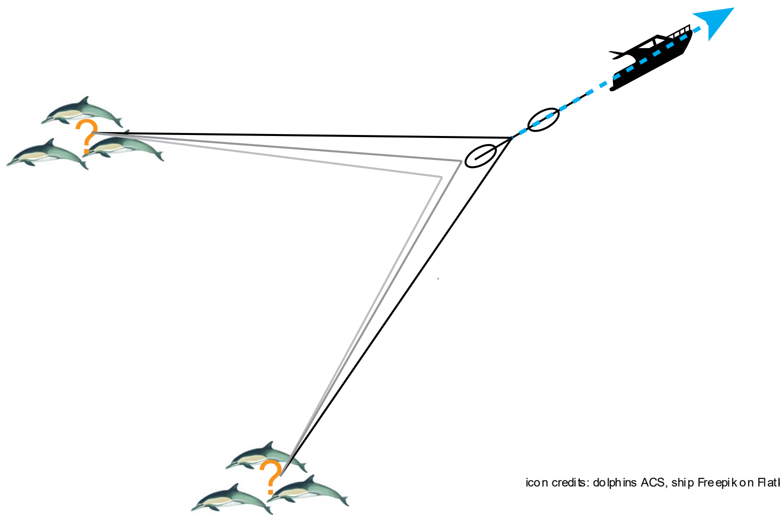

Figure 3- Positions from multiple bearing angles.

In this scenario, we will assume that each localized position produced by a set of bearings will be represented as offsets in meters from a fixed latitude and longitude (Subtype Engineering, Name Cartesian with a WGS84 Anchor). Other representations of coordinates are of course possible, such as producing Cartesian coordinates relative to the instrument, angular coordinates, UTM coordinates, or converting these positions to WGS84 longitudes and latitudes. When localizing relative to a mobile instrument it is worth thinking about how these data are likely to be used. In many cases, it will simplify future analyses if the data are converted to positions relative to a fixed datum.

The Id, Description, DataSource, Algorithm, QualityAssurance, ResponsibleParty, and UserId structures and fields would all be populated in a similar manner as has been described for other examples.

The Effort structure associated with these localizations will have several differences from the one that we used to specify the bearing effort of the previous example. The CoordinateReferenceSystem Subtype will be set to "Engineering," and the Name to "Cartesian." In the ReferenceFrame, we set the Anchor to "WGS84," and Latitude and Longitude to the fixed location from which we will be calculating offsets in meters.

Dimension will be set to 2 and the LocalizationType to "Point". Crossed bearing methods assume that animals have not moved significantly between bearings estimates, and consequently the time of localization can be some function of the timestamps associated with the bearings. The effort TimeReference will be recorded as "relative." How the timestamp is set should be inferable from information recorded in the Algorithm record.

The localizations derived from the bearings would be stored in Localization records. Each Localization may have an optional Event text associated with it which must be unique to this set of localizations and can allow other derived data (e.g., tracks) to reference specific localizations. TimeStamp indicates the date and time assigned to the localization. If we wish to link the position to the bearing records used to produce them, this would be stored in the EventReference field of the Reference record.

This is followed by a Cartesian record which contains offsets to WGS84 Anchor defined in the effort section in meters (X_m, Y_m). As in the bearing example, CoordinateError permits recording the standard error of this coordinate. The final field of the localization record is the optional QualityAssurance which can be set to unverified, valid, or invalid.

5.4 Point localizations on a Navy Range

The US Navy maintains several wide-aperture hydrophone ranges for training that are also used to study marine mammals. These hydrophones could be treated as a single deployment, an ensemble with a large number of hydrophones, or sets of smaller ensembles.

In this scenario, we will make the following assumptions:

- Deployment records were created for each hydrophone.
- An ensemble of 12 hydrophones has been created. See section 6 (p. 30) for an example scenario.
- Detection algorithms were previously executed, and the resultant detections are already stored. Each detection has an event identifier that is unique to the set of detections.
- Localizations are measured using latitude and longitude (WGS84) and no elevation data are produced.

The first step is to assign a unique Id string to the set of localizations. The following mandatory fields would then be set: DataSource, Algorithm, and UserId. The first three fields all have the same semantics as for detections and describe what data are being processed, how they were processed, and the user that processed these detections.

Effort for this scenario would include date and time timestamps denoting the start and end of analysis. This would be followed by a specification of the CoordinateReferenceSystem, with Subtype=Geographic, Name=WGS84. The remainder of the Effort record would indicate Dimension=2 (as no elevation/depth is estimated), LocalizationType=Point as we are producing locations and TimeReference=relative indicating that localization timestamps are some function of the received times on the various channels.

The Localizations element contains a list of Localization elements. Each Localization element may contain an optional event identifier (Event) that is unique to the set of localizations. Timestamp denotes the date and time of the localization point estimate and may correspond to arrival time on any of the channels on which the signal was detected or some function of those times. To preserve a history of the detections used to produce the localization (optional), the EventReference field in the References record would contain a list of detections indicating the Id of the detections group from which they came and the Event identifier associated with the detection..

Each Localization contains one WGS84 record. Its Coordinate record will have the Longitude and Latitude ([-90, 90] °N and [0, 360) °E). Note that while we store longitudes over a standard interval to make it easier to design algorithms that search for values, it may be advantageous to work with coordinates on a different interval. This is especially true when working near longitudes of 0°, and it is trivial to convert retrieved data to a more convenient format with the following formula MOD(latitude + 180, 360) – 180 where MOD(a,b) is the modulo, or remainder of a / b.

5.5 TRACKS ON A NAVY RANGE

Only minor changes are needed to extend the previous scenario to include information about animal tracks. The major changes from the previous scenario are:

- In the CoordinateReferenceSystem record, LocalizationType is set to Track.
- Each localization record consists of a Track subrecord.

Track localization records a Timestamp and a location type which in this scenario is WGS84. The contents of Timestamp and WGS different from the point estimates in that they have vector values denoting times, locations, and errors along the track. An optional Bounds record can provide a bounding box for WGS84, UTM, and Cartesian coordiantes. It contains records NorthWest and SouthEast with X, Y, and Z values in the units of the current coordinate reference system. X and Y are the northwest and southeast values in WGS84 and UTM, and the left-uppermost and right-lowermost values for Euclidean measurements. Regardless of the coordinate reference system Z is the minimal value for the northwest entry and the maximal for southeast. Providing a bounding box can enable the construction of efficient indexing schemes for fast retrieval.

6 ENSEMBLES

Ensemble records are used to denote a collection of temporally overlapped deployments that are being used together. Sample applications include tasks such as beam-forming to improve signal strength or localization. In many cases, these tasks are performed from recordings that were part of separate instrument deployments. Ensemble records provide identifiers that can be used to group instruments together when this situation arises.

6.1 ENSEMBLES FOR DETECTION AND LOCALIZATION

Guazzo *et al.* (2017) conducted a study of gray whale (*Eschrichtius robustus*) migration by localizing M3 calls which typically have a bandwidth of approximately 20-200 Hz. Low frequencies require larger spacing between transducers to conduct localization. They deployed 4 passive acoustic recorders separated by approximately 1-2 km from one another south of California's Point Lobos State Marine Reserve. Deployment records would be created for each of these instruments as outlined above. An ensemble record links these deployments together as a common entity. Note that in cases where all transducers are connected to a single instrument and thus contained in a single deployment, there is no need to create an ensemble record.

Ensemble records are identified by a unique string called the Id. Each deployment associated with an ensemble is called a Unit. Each of the 4 units in this scenario would record a DeploymentId containing the Id field associated with each of the deployments, and an integer UnitId that would allow the deployments to be referenced by ensemble name and unit. Together, the ensemble's Id and UnitId fields allow other records to easily refer to a specific deployment within the ensemble.

A geographic anchor for the ensemble is defined in the ZeroPosition field. It could be the geographic center of the units, the position of one of the deployments, or an arbitrary point. The selected location must be specified in a standard coordinate system such as WGS84 or UTM. An optional ElevationInstrument_m permits a vertical measurement relative to average sea level to be associated with the center point.

In this sample scenario, Guazzo *et al.* (2017) used the generalized power law detector (Helble *et al.*, 2012) to process all detections. These detections could have been recorded as detections with respect to individual deployments, or the ensemble. To record detections with respect to the ensemble, each detection would note the ensemble Id and UnitId during the detection process.

During the localization phase, one could choose to record identifiers for specific detection events used for each localization. This permits the retention of the knowledge about which detections on which instruments were used to localize a sound source. The ZeroPosition would allow recording of locations relative to a specific point if the localization effort produced relative distance measurements.

While we do not provide a separate scenario for a beam-forming application, the approach would be similar. The detection process could record information about the units of the ensemble that were used to enhance the signal.

7 LIST OF COMMON RECORDS AND FIELDS

1. BespokeData – This field can be used to reference non-standard data sources. For example, one might wish to store additional data related to localizations such as correlation scores, etc. or add summary information about hundreds of millions of echolocation clicks while retaining details about peak spectra, received levels, etc. The BepokeData permits this. Fields:
   - Abstract – Textual description of the data that are being stored.
   - Data – A list of subfields indicating where the data may be found. Contains repeated instances of:
     - URI – Uniform Resource Indicator (e.g., web address, digital object identifier) that indicates where the data may be found. It is possible to us descriptive text here instead of a URI, but retrieval may be more complicated. Note that database systems implementing this standard are not required to handle the archiving of these data.
     - Comment – Optional text describing this specific data resources, mainly useful when there are multiple URIs.
   - UserDefined – A list of user defined fields and values, may be nested. Use Data/URI for large amounts of data.

2. Description – A human readable description of metadata. Most descriptions are optional, but are recommended as they let the user of the data understand what was intended at a high level. Descriptions may contain:
   - Objectives – What are the objectives of the effort?
   - Abstract – An overview of what was done.
   - Method – A high level description of the work that was done, such as a quick summary and citation of a paper that describes an algorithm/method.
3. Id – A collection of letters, numbers, and punctuation (a string) that uniquely identifies a set of metadata within a group of similar items. For example, an Id allows us to uniquely identify one deployment amongst all other deployments.
4. MetadataInfo – For many large projects, knowing who is responsible for metadata records is critical. A common facet of all of the types of information discussed in this document is an optional MetadataInfo element. It contains components of the ISO 19115-1 ResponsibleParty specification: a person's name, their organization, their position name, and one or more sets of contact information (phone, address, online resources, hours of service, and contact instructions). The ResponsibleParty's onlineResources element does not conform completely with ISO 19115-1. It also specifies a timestamp indicated when the information was last updated[1] (Date) and how often updates are planned (UpdateFrequency): as-needed, unplanned, or yearly.
5. ResponsibleParty – A subset of the ISO 19115-1 standard, this includes text strings for individualName , organizationName, positionName, and a contactInfo element. The contactInfo consists of a phone element with zero or more voice and facisimile numbers, address information with zero or more deliveryPoint strings for multiline addresses, a city, administrativeArea (e.g. state), postalCode, country, and electronicMailAddress. Optional elements permit designation of service hours (hoursOfService) and instructions on contacting the party (contactInstructions).
6. Timestamp – A day and time. All dates and times are stored in universal coordinated time (UTC) using a subset of the ISO 8601 standard: YYYY-MM-DDTHH:MM:SSZ, where Z indicates UTCS time, e.g. 1940-02-12T14:10:00Z. Optional time zone modifiers are permitted to store data in local time with an offset, but all data are stored in UTC.

8 GLOSSARY

ISO 19111:2019 – Referencing by coordinates. Portions of this standard have been adopted in the localization schemata.

ISO 19115-1:2014 – Geographic information, Metadata. We adopt portions of this standard.

ISO 8601 – Date and time format. Timestamps are of the format YYYY-MM-DDTHH:MM:SSZ where Z indicates universal coordinated time (UTC). There are provisions for local time which are represented by a + or – offset of HH:MM following the Z, but within this work all timestamps are stored in UTC.

[1] It is unclear whether this is supposed to be the last update for the first time it is added. We will need to check with an ISO 19115 expert.

9 References

Berners-Lee, T., Fielding, R., Day Software, Masinter, L., and Adobe Systems (2005). "Uniform Resource Identifier (URI): Generic Syntax," Accessed March 29, 2024. https://datatracker.ietf.org/doc/html/rfc3986.

Gassmann, M., Elizabeth Henderson, E., Wiggins, S. M., Roch, M. A., and Hildebrand, J. A. (2013). "Offshore killer whale tracking using multiple hydrophone arrays," J Acoust Soc Am 134(5). 3513-3521, doi:10.1121/1.4824162.

Gillespie, D., Gordon, J., McHugh, R., McLaren, D., Mellinger, D. K., Redmond, P., Thode, A., Trinder, P., and Deng, X.-Y. (2008). "PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans," in *Proc. Inst. Acoustics* (Inst. Acoust., Southampton, UK, October 14-15), pp. 54-62.

Guazzo, R. A., Helble, T. A., D'Spain, G. L., Weller, D. W., Wiggins, S. M., and Hildebrand, J. A. (2017). "Migratory behavior of eastern North Pacific gray whales tracked using a hydrophone array," PLoS One 12(10), pp. e0185585, doi:10.1371/journal.pone.0185585.

Helble, T., Ierley, G. R., D'Spain, G. L., Roch, M. A., and Hildebrand, J. A. (2012). "A generalized power-law detection algorithm for humpback whale vocalizations," J Acoust Soc Am 131(4). 2682-2699.

International Standards Organization (2003). "Geographic Information - Metadata," ISO 19115:2003, International Standards Organization, Geneva, pp. 140.

ITIS Organization (2014). "Integrated Taxonomic Information System (ITIS)," Accessed July 4, 2014. http://www.itis.gov.

Klinck, H., Winiarski, D., Mack, R. C., Tessaglia-Hymes, C. T., Ponirakis, D. W., Dugan, P. J., Jones, C., and Matsumoto, H. (2020). "The Rockhopper: a compact and extensible marine autonomous passive acoustic recording system," in *Global Oceans 2020: Singapore – U.S. Gulf Coast* (pp. 1-7, doi:10.1109/ieeeconf38699.2020.9388970.

Mellinger, D. K. (2001). "Ishmael 1.0 User's Guide," OAR-PMEL-120, NOAA PMEL, Seattle, pp. 30.

Open Geospatial Consortium (2023). "OGC Standards and Supporting Documents," Accessed June 1, 2023. http://www.opengeospatial.org/standards.

Soldevilla, M. S., Wiggins, S. M., and Hildebrand, J. A. (2010). "Spatial and temporal patterns of Risso's dolphin echolocation in the Southern California Bight," J Acoust Soc Am 127(1). 124-132, doi:10.1121/1.3257586.

Wiggins, S. M., and Hildebrand, J. A. (2007). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring," in *Intl Symp Underwater Tech* (Tokyo, Japan), pp. 551-557, doi:10.1109/UT.2007.370760

Wiggins, S. M., and Morris, M. A. (2019). "SoundTrap ST500 Calibration at the Transducer Evaluation Center (TRANSDEC)," in *Marine Physical Laboratory*, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, San Diego, CA.