Standards for Acoustic Metadata for Passive Acoustic Monitoring

DRAFT STANDARD

Subcommittee 1, working group 7:

Table of Contents

Foreword

The ANSI standard for acoustic metadata provides a set of standard vocabulary for preserving information derived from acoustic recordings with a focus on serving the bioacoustics community. The standard enables preservation of data related to the characteristics of instrumentation used to record acoustic data, which methods were used to detect and characterize sounds of interest within these recordings as well as what was found, and which methods were used to determine directional or positional information of the sound sources.

A critical component of this standard that is not well-represented in other standards related to sensors or to the more generalized problem of studying animals is the detailed tracking of effort, methods, and results used in the characterization of sound. The standard provides extensive capabilities to denote both details and variations in recording and analysis effort. In addition to retaining information needed to make studies reproducible such as noting details about an algorithm that was used in analysis such as the implementation version and tunable parameters, it is critical to understand to which data were the algorithm applied. This is not tracked in many standards. Without knowing the temporospatial extent over which an analysis method is applied, valid inference is impossible. As an example, if one only knows that a set of blue whale (*Balaenoptera musculus*) calls were detected in June 2024 but did not know over what period the analysis was conducted, one would be unable to distinguish between blue whales being present only in June, or being present year round with analysis only conducted over a subset of the data.

This standard comprises a part of a group of definitions, standards, and specifications for use in acoustics. It was developed and approved by Accredited Standards Committee S3/SC1 Animal Bioacoustics, under its approved operating procedures. Those procedures have been accredited by the American National Standards Institute (ANSI). The Scope of Accredited Standards Committee S3/SC1 is as follows:

> *Standards, specifications, methods of measurement and test, instrumentation, and terminology in the field of psychological and physiological acoustics, including aspects of general acoustics which pertain to biological safety, tolerance, and comfort of non-human animals, including both risk to individual animals and to the long-term viability of populations. Animals to be covered may potentially include commercially grown food animals; animals harvested for food in the wild; pets; laboratory animals; exotic species in zoos, oceanaria or aquariums; or free-ranging wild animals.*

This standard is not comparable to any existing ISO/IEC Standard.

This standard is not a modified nationally adopted international standard.

This standard includes 2 Annexes. Annex A is normative and is considered to be a part of this standard. Annex B is informative and is not considered part of this standard.

At the time this Standard was submitted to Accredited Standards Committee S3/SC1 Animal Bioacoustics, for approval, the membership was as follows:

Marie A. Roch, *Chair*
Shane Guan, *Vice-Chair*

Nancy Blair-DeLeon, *Secretary*

Acoustical Society of America .... Aaron M. Thode
.... Paul D. Schomer (Alt.)

G.R.A.S. Sound & Vibration .... Robert O'Neil

National Park Service .... Kurt Fristrup
.... Megan McKenna (Alt.)

Ocean Conservation Research .... Michael Stocker

U.S. Army Aeromedical Research Laboratory .... William A. Ahroon

U.S. Army CERL .... David K. Delaney
.... Michael J. White (Alt.)

U.S. Navy Marine Mammal Program .... James J. Finneran
.... Elizabeth Henderson (Alt.)

University of Cincinnati Animal Audiology Clinic/Bioacoustics Lab .... Peter M. Scheifele
.... David K. Brown (Alt.)

Individual Experts of Standards Subcommittee S3/SC1, Animal Bioacoustics, were:

Ann E. Bowles	Kurt Fristrup	Aaron Thode
David K. Delaney	Darlene Ketten	
James J. Finneran	Colleen Reichmuth	

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:

Marie A. Roch, Chair, San Diego State University, San Diego, CA, USA
Simone Baumann-Pickering – Scripps Institution of Oceanography, Univ. California, San Diego, CA, USA
Danielle Cholewiak, National Oceanographic and Atmospheric Administration Northeast Fisheries Science Center
Douglas Gillespie, St. Andrews University, St. Andrews, Scotland
Shane Guan, Bureau of Ocean Energy Management, Sterling, VA, USA
Jasper Kanes, Ocean Networks Canada, Victoria, BC, Canada
Katherine H. Kim, Greeneridge Sciences, Inc, Santa Barbara, CA, USA
Holger Klinck, Cornell University, Ithaca, NY, USA
Xavier Mouy, National Oceanographic and Atmospheric Administration Northeast Fisheries Science Center (formerly of JASCO Applied Sciences, Victoria, BC, Canada)
Ana Širović – Norwegian University of Science and Technology, Trondheim, Norway
Aaron Thode – Scripps Institution of Oceanography, Univ. California, San Diego, CA, USA
Carrie Wall – NOAA National Centers for Environmental Information, Boulder, CO

Suggestions for improvements to this standard will be welcomed. They should be sent to Accredited Standards Committee Accredited Standards Committee S3/SC1, Animal Bioacoustics, in care of the Standards Secretariat of the Acoustical Society of America, 1305 Walt Whitman Road, Suite 300, Melville, New York 11747. Telephone: 631-390-0215; FAX: 631-923-2875; E-mail: standards@acousticalsociety.org.

Introduction

Sound plays an important role in nature, with many animals relying on sound for signaling between conspecifics and across species (Bradbury and Vehrencamp, 1998; Au and Hastings, 2008). As a consequence of this, there are extensive efforts to analyze passive acoustic data collected by recorders in both marine and terrestrial settings (see Stowell, 2022 for a review) as well as to characterize the growing contributions of anthropogenic activities to the acoustic environment (e.g., Hildebrand, 2009; Shannon *et al.*, 2016). Whether conducted by human analysts or by machine learning and signal processing algorithms, these efforts generate derived data products describing information about the contents of the acoustic recordings. In general, the data products generated by these efforts are stored using group or algorithm specific methods that make the analysis of long-term trends across wide geographic regions difficult.

Scope

This standard represents an effort to codify the information related to the characterization of sound produced by or impacting animal populations. These bioacoustic data should be stored with the following goals: to make comparisons between different efforts possible, and to enable the creation of data repositories that can provide long-term storage and sharing. Intended users of this standard are scientists using passive acoustic monitoring to study behavior and ecology, companies and governments monitoring wildlife as part of a mitigation or census strategy, and policy makers who wish to make scientifically informed decisions based upon the results of acoustic studies.

The standard differs from existing standards in terms of purpose and scope. Several standards are related to geographic information. The ISO-19115 standard (International Standards Organization, 2003) was designed for describing equipment and measurement processes within a geographic context, and was revised in 2019. Similarly, the Open Geospatial Consortium maintains a wide series of standards related to geographic information, measurement processes, and physical metadata such as meteorlogical measurements (Open Geospatial Consortium, 2023). When possible, we adopt elements of these standards, but both families of standards were not designed for biological contexts and are missing elements critical to bioacoustics measurement such as methods of detailing analysis effort and description of biological sounds. The Integrated Ocean Observing System Association (IOOS, https://ioosassocation.org) is another effort to design standard vocabulary within an ocean context. Similar to the geographic standards above, IOOS is excellent for describing equipment deployments and physical measurements, including describing the measurement of acoustic data (Guan *et al.*, 2014), but current efforts for biological data (IOOS Biology, 2018) are also insufficient for bioacoustics data.

A number of standards are designed to record information about biological systems. Darwin Core (Wieczorek *et al.*, 2012) is designed to provide geospatial descriptions of animals. Historically, Darwin Core has not addressed issues such as recording systematic effort used to detect animals, but more recent versions of Darwin Core (Darwin Core Maintenance Group, 2021) have introduced terms such as event and sampling effort which capture information about sampling strategies and methods using an uncontrolled vocabulary which complicates search and the use of multiple efforts in meta-data studies. Vocabulary have also been introduced to denote sounds produced by animals, but not in a systematic way. OBIS-SEAMAP (Halpin *et al.*, 2009) maintains a standard way of recording geospatial animal observations, and an extension by Fujioka *et al.* (2014) permits the integration of acoustic observations, although this is done at a summary level rather than permitting the annotation of individual sounds.

The lack of bioacoustic-specific standards led to efforts to establish a community standard vocabulary to describe data derived from acoustic recordings (Roch *et al.*, 2013; Roch *et al.*, 2016). This standard generalizes and extends the data model proposed in this earlier work.

Standards documents can frequently be difficult to read as they list how something should be done while giving very little insight as to why. Annex A for this standard demonstrates how the standards support acoustic collection and analysis through a series of scenarios. While Annex A informs how the standard is used, it is not a substitute for the standard itself.

This standard is divided into several sections, each addressing different aspects of acoustic data collection and analysis. Some concepts used to describe data collection and analysis occur in multiple sections. When this is the case, we note that the concept is used in multiple places and provide the specification in a separate section (p. 26) on common data. We begin by addressing deployments of acoustic collection devices, then specify how calibrations of these instruments should be recorded. A section on ensembles discusses how disparate deployments can be logically joined into a virtual unit for activities such as beam forming or localization.

The next two sections address information that is derived from these deployments. In the first, we discuss detections which can range from simple presence/absence metrics to annotations of individual calls or recording soundscape metrics. We then discuss location information, including concepts such as bearings, positional information, and tracks.

In this standard, we provide a set of standard names and data types. The standard does not detail how these data are stored. There are many implementation choices about how to store this specification, with a few of the options consisting of relational databases, key-store database, network structure databases, etc. Implementation is a vendor-specific choice and there are pros and cons of each organizational strategy. A reference implementation of the standard is provided in the open-source Tethys acoustic metadata server (https://tethys.sdsu.edu) which uses a network structure database that stores extended markup language (XML) data. Additional projects that conform by and large to the standard include the US National Centers for Environmental Data's Passive Acoustic Data archive (https://www.ncei.noaa.gov/products/passive-acoustic-data) and the National Oceanic and Atmospheric Administration's Passive Acoustic Cetacean Map (https://www.fisheries.noaa.gov/resource/data/passive-acoustic-cetacean-map).

Creating a standard that meets the needs of all users is a difficult task, as there is not a single correct way to specify things that is appropriate for every situation. Consequently, the committee has focused on determining which data need to be retained for a wide variety of bioacoustics applications. The committee has tried to strike a balance between providing something that is useful to the vast majority of bioacousticians and policy makers without burdening the user with an overly cumbersome set of requirements. While we provide a standard set of names and data types, we recognize that people with existing methods of representing their metadata may have chosen to represent things differently or have different names for similar concepts. In these cases, if users do not wish to convert to the new names, determining how things map from the existing system to the standard is frequently sufficient for building a data mapping scheme that would permit sharing of standardized data. As an example of this, the standard encourages the use of the term elevation to describe instrument deployments, with elevation specified in meters above (+) or below (-) average sea level. In marine environments, it is more common to use the term depth, and a system that records depth can certainly continue using that term. The standard, which is written for both marine and terrestrial systems, encourages one to use elevation, but a system that records depth with the understanding that positive depth values can be mapped to negative elevation values would not prevent information sharing with other systems that are compliant with the standard.

Within each section, items are grouped thematically and marked as mandatory (M), optional (O), or conditional (C). Some items may be repeated, and we mark such items with a + followed by the number of times that something may be repeated. +∞ indicates that the item may be repeated an arbitrary number of times. Items are frequently grouped into records, and records are marked as mandatory or optional. Items within an optional record may still be marked as mandatory, indicating that they are required when that record is used. As an example, quality control information about a recording may be optional, but when it is recorded, there will be some mandatory quality control fields. Items are marked as floating point ($\mathbb{R}$), integer ($\mathbb{Z}$) non-negative integers ($\mathbb{Z}^+$), text (ᴀA), position (🌎), and date & time (🕑). When items are grouped, the group name will not show a data type. The list of symbols is summarized in Table 1.

Table 1 – Annotation symbols used to describe data fields.

Symbol	Explanation
C	Conditional field – Field required when other fields have certain values
M	Mandatory field

O	Optional field
+N	Field may be repeated N times, +∞ indicates an unlimited number
$\mathbb{R}$	Real number (floating point)
$\mathbb{Z}$	Integer
$\mathbb{Z}^+$	Non-negative integer (0, 1, 2, …)
$\overrightarrow{\mathbb{R}}$, $\overrightarrow{\mathbb{Z}}$, $\overrightarrow{🕑}$	A vector of values for a specified type, e.g., list of integers: $\overrightarrow{\mathbb{Z}}$
ᴀA	Text
🕑	Timestamp – date and time, always stored in universal coordinated time (UTC)
🌎	Position in a standard coordinate system such as world geodetic system 84 (WGS84 latitude and longitude) or universal transverse Mercator (UTM)

While the standard is implementation agnostic, there are requirements that specific types of data meet minimal requirements. An example of this is the timestamp field which is mandated to be able to record date and time to at least the microsecond. Requirements are detailed in annex A, Data Types, on page 32. In order to promote interoperability between different implementations of this standard, we also recommend that software be created to permit transformation from the data in specific implementations to a common representation. The current standard recommends that data be transformed to XML documents that conform to XML schema specification (Walmsley, 2002) associated with this standard (see https://tethys.sdsu.edu/standards). It is possible to transform these into other standardized information interchange formats such as Javascript object notation (JSON Schema, https://json-schema.org/), and other forms of information interchange may be used.

Normative References

The following referenced documents are indispensable for the application of this standard. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

Deployments

A deployment record should be generated when an acoustic recorder with one or more transducer (microphone/hydrophone) elements is placed in the field. Deployments take many forms, ranging from a stationary recorder mounted on a tree to an array of hydrophones towed behind a boat. It is assumed that the transducers in a deployment are time synchronized and that the corresponding channels within the deployment are time aligned. If a deployment uses multiple instruments (e.g., multiple data loggers at different depths on a mooring line that each have their own clocks that drive sampling), it is generally preferable to consider these as separate deployments.

Each deployment shall record the following information:

1. Id (ᴀA M) – Text that uniquely identifies the instrument deployment.
2. Description (O) – A record containing optional prose about the objectives, abstract, and method associated with this deployment. See Common Data Fields (item 2, p. 27) for further details.

3. Project (ᴀA M) – Text that indicates the project with which this deployment is associated. Typically, this is related to a geographic region, funding source, etc.
4. DeploymentId ($\mathbb{Z}$ M) – Used to help distinguish groups of deployments set to a value of the user's discretion. The intended use is to denote the Nth deployment of an instrument in a sequence, such as the Nth deployment of a stationary instrument at a specific site or the Nth tow of an instrument along a cruise track. In situations where this field does not provide value, a value of 0 is recommended.
5. DeploymentAlias (ᴀA O) – Text used to provide an alternative description of a deployment.
6. Site (ᴀA O) – Text used to provide a name for a deployment location. Examples include place descriptions such as "Tanner Banks" or project-based index schemes such as providing letters/numbers to designate specific sites within a Project.
7. SiteAliases (O) – A list of alternative Site (ᴀA $+\infty$) subrecords. Each Site subrecord consists of text.
8. Cruise (ᴀA O) – Text providing the name of the deployment vehicle associated with the deployment.
9. Platform (ᴀA M) – Text indicating the platform on which the instrument was deployed. These strings are not restricted to standardized values. Examples include: mooring, animal-borne tag, towline, etc. For marine deployments, it is recommended to use the IOOS platform vocabulary (Guan *et al.*, 2014) when possible:
    - glider – underwater glider
    - drifting_buoy – A buoy that drifts with the current
    - moored_buoy – A buoy that is secured with a mooring line.
    - bottom_mounted – An instrument that is part of a sea-floor package.
10. Region (ᴀA O) – Text providing the name of the geographic region where the instrument was deployed. This is useful when Project is associated with something other than a geographic region.
11. Instrument (M) – Instrument consists of 3 subfields:
    a. Type (ᴀA M) – Text with name of instrument family, e.g., HARP, EAR, Popup, DMON, Rock Hopper, LS1, SoundTrap, etc.
    b. InstrumentId (ᴀA M) – Text providing a unique identifier for an instrument, typically a serial number.
    c. GeometryType (ᴀA O) – Sensor configuration type, must be one of two text values:
    - rigid – relative geometry of transducers is fixed
    - cabled – relative geometry may be expected to deform

    The position of sensors relative to the instrument is defined later in the sensors section.

12. SamplingDetails (M) – Information about the instrument's sampling plan. SamplingDetails are collected for one or more Channel (M, $+\infty$) records. Each Channel record contains:
    a. ChannelNumber ($\mathbb{Z}^{+}$ M) – Index of an acoustic data stream within a recording. There is no mandatory channel numbering scheme, but we assume that most recordings will have channels 0 to N-1 or 1 to N where N is the number of channels recorded. These should correspond to the channel streams within standardized storage formats such as resource information interchange format's WAVE (Arms *et al.*, 2023) files.

b. SensorNumber ($\mathbb{Z}^+$ M) – Index of a physical sensor that was used to collect the channel data. Physical sensor characteristics such as hydrophone/microphone models are defined in the AudioSensors element that is described later.
c. Start (🕑 M) –Start of recording effort.
d. End (🕑 M) – End of recording effort. Scheduled recording and unplanned data gaps can be handled with the DutyCycle (12.h) and Quality (13.c) fields respectively. In general, if an instrument is recovered and later deployed again such as a daily deployment of an array on a cruise, separate deployment records should be generated.
e. EventTrigger (O) – Some instruments only record when specific events occur, such as during the detection of certain types of sounds. When present, the record contains up to two items:
    1. Description (O) – A general description of the triggering algorithm. See common fields Common Data Fields (item 2, p. 27) for further details.
    2. Algorithm (M) – A description of the triggering algorithm. See Common Data Fields (item 3, p. 27) for further details. It is recommended that the Algorithm/Parameters record be used to capture information related to trigger conditions and what is saved.
f. Sampling (M) – Description of how data were sampled. A list of sampling Regimen (+∞) fields shows the start of each sampling configuration. Each Regimen remains active until the next Regimen field or the end of the recording. Each sampling regimen contains:
    1. Timestamp (🕑 M) – Date and time the sampling plan started.
    2. SampleRate_Hz ($\mathbb{R}$ M) A floating point number describing the sample rate in Hz.
    3. SampleBits ($\mathbb{Z}^+$ M) - Number of quantization bits used to represent each sample.
g. Gain (O) – Used to denote deviations from the standard gain that would be achieved by a calibration at standard settings. A list of gain Regimen (+∞) fields show the start of each gain configuration (similar to sampling regimens). Each gain Regimen contains:
    1. Timestamp (🕑 M) – Date and time of the gain change.

    A specification of gain with either:

    2. Gain_dB ($\mathbb{R}$ M) Gain boost in decibels.

    or:

    3. Gain_rel ($\mathbb{R}$ C) – Uncalibrated gain boost such as a number on a dial. As most scientific recordings are calibrated, use of this field is expected to be rare.
h. DutyCycle (O) – Denotes periodic recordings and should not be used when recording is continuous. A list of duty cycle Regimen (+∞) fields show the start of each duty cycle configuration (similar to sampling regimens). Each duty cycle Regimen contains:
    1. Timestamp (🕑 M) – Start of the duty-cycled schedule.
    2. RecordingDuration_s ($\mathbb{R}$ M) – Number of seconds of continuous recording.

        3. Offset_s ($\mathbb{R}$ O) – Number of seconds past the specified Timestamp (12.h.1) before the recording starts. If omitted, defaults to an offset of 0.
        4. RecordingInterval_s ($\mathbb{R}$ M) – The number of seconds between starts of consecutive periods.

        Constraint: RecordingDuration_s + Offset_s ≤ RecordingInterval_s

        Example: RecordingDuration_s = 300, RecordingInterval_s = 600 implies recording the first five minutes of every ten minutes starting at the specified timestamp.

13. QualityAssurance (O) – Information related to any quality assurance processes that have been conducted on the data. Contains the following subfields:
    a. Description (ᴀA O) – Description of the process.
    b. ResponsibleParty (O) – Denotes the person or entity responsible for the quality assurance record. See Common Data Elements for details.
    c. Quality (M, +∞) – Each Quality record contains:
        1. Start (🕑⃗ M) – Record of start date and time.
        2. End (🕑⃗ M) – Record end date and time.
        3. Category (ᴀA M) – One of the following text values:
   - unverified – Indicates that the data have not been verified.
   - good – Data have been inspected and are of good quality.
   - compromised – Data may be usable, but have anomalies.
   - unusable – The data are not valid.
        4. FrequencyRange (O) – Shows the frequency range over which the quality annotation applies.
   - Low_Hz ($\mathbb{R}$ M) – Lower limit in Hz.
   - High_Hz ($\mathbb{R}$ M) – Upper limit in Hz.

        Constraint: Low_Hz ≤ High_Hz

        5. Channel ($\mathbb{Z}^+$ O) – Integer indicating to which channel the annotation applies. If absent, it is assumed to apply to all channels.
        6. Comment (ᴀA O) – Prose indicating the nature of the damage. Example: "Low frequencies unusable for biotic detections due to tidal flow."
14. Data (M) – Descriptions of where to find data. Data values can be included for some low-volume non-audio sources.
    a. Audio (M) – Information on how to access the audio data. Contains the following elements
        1. URI (ᴀA M) – Indicates where audio data may be accessed. If the text conforms to a uniform resource indicator (URI, Berners-Lee, 2005), that URI may be used to access the resource. Example URIs include digital object identifiers, web pages, file transfer protocol servers, etc. This is a text field and if a URI is unavailable, a description can be used instead; software can easily distinguish between valid URI data and text such as "filing cabinet 5."
        2. Processed (ᴀA O) – Text about where to access derived data products such as decimated or filtered data. May contain a URI.

3. Raw (ᴀA O) – Text or URI indicating where to access original data from the instrument if it has been archived separately from the data specified by the URI field.

b. Tracks (O) – This field is used to denote geospatial track information for mobile instruments.
    1. SpeedUnit (ᴀA C) – Required if tracks contain measurements of speed. If present, must be one of the following:
        kn – nautical miles per hour
        km/h – kilometers per hour
        m/s – meters per second
    2. Track (O, +∞) – Information on a specific track

- TrackId (ℝ O) – Floating point number specifying an identifier for a track. This is expected to be unique within the deployment. If tracks are not numbered, this element is not required.
- Point (M, +∞) – Information about the current track point. Only the Timestamp child field is mandatory as we do not expect different sensors to be sampled at the same rate.
    - Timestamp (🕑 M) – Date and time at which point measurement was made.
    - Position (🌎 O) – Position in a standard coordinate system.
    - Heading_DegN (ℝ O) – Counterclockwise angle (degrees) between the front (or bow) of the instrument and true North.
    - CouseOverGround_DegN – Counterclockwise angle between the instrument's current motion vector and either magnetic or true North (see north field).
    - CourseOverGround_North (ᴀA C) – Reference vector for course over ground measurements. Required when CourseOverGround_DegN is present. Must be one of:
        true – Reference is true north
        magnetic – Reference is magnetic north
    - Speed (ℝ O) – Speed relative to the medium (e.g., air, water, which itself may be in motion) in which the deployment package is situated, in units specified by SpeedUnit (14.b.1, required when Speed is present).
    - SpeedOverGround (ℝ O) – Speed relative to the ground, in units specified by SpeedUnit (14.b.1, required when SpeedOverGround is present).
    - Pitch_deg (ℝ O) – Instrument pitch in degrees, restricted to: [0, 360)
    - Roll_deg (ℝ O) – Instrument roll in degrees, restricted to: [0, 360)
    - Elevation_m (ℝ O) – Instrument elevation in meters.
    - GroundElevation_m (ℝ O) – Elevation of ground or sea floor in meters relative to sea level.

3. TrackEffort (O) – Used to distinguish between systematic track effort and derivations from planned tracks.
    - OnPath (M, +∞) – Description of a portion of Track that is considered to be part of the planned survey effort. Contains the following subfields:
        i. TrackId (ℝ O) – Identifies a TrackId within Track.
        ii. Start (🕑 M) – Start of OnPath track effort.
        iii. End (🕑 M) – End of OffPath track effort.
        iv. Comment (🗛 O) – User notes.
        v. FocalArea (🗛 O) – Name of a focal area of interest associated with this entry, such as a sanctuary or preserve.
    - OffPath (O, +∞) – Description of portion of Track that deviates from planned survey effort. Contains identical fields to OnPath with the same semantics (see above):
        i. Trackline (ℝ O)
        ii. Start (🕑 M)
        iii. End (🕑 M)
        iv. Comment (🗛 O)
        v. FocalArea (🗛 O)
    - URI (🗛 O +∞) – Uniform resource indicator (web address) detailing where additional sources of track data may be found.

15. DeploymentDetails (M) – Information about initial deployment of an instrument. Contains the following fields:
    a. Position (🌎 M) – Position in a standard coordinate system.
    b. ElevationInstrument_m (ℝ O) – Elevation of the instrument relative to sea level in meters. Negative numbers indicate meters below sea level.
    c. DepthInstrument_m (ℝ O) – Depth below the sea or ground surface in meters. This is usually the negative of ElevationInstrument_m, in which we case it is recommended to use ElevationInstrument_m and not use this field. Some applications, such as recording in mountain cave systems and alpine lakes may require a depth reading that is relative to a local surface that is not at sea level in which case DepthInstrument_m should be used.
    d. Elevation_m (ℝ O) – Elevation of ground or sea bed relative to sea level.
    e. Timestamp (🕑 M) – Date/time at which instrument was deployed.
    f. AudioTimeStamp (🕑 M) – Date/time at which the recording schedule starts.
    g. Vessel (🗛 O) – Name of deployment vehicle
    h. Person (O) or ResponsibleParty (O) – Contact information may be specified using either Person or ResponsibleParty; see the section on common data fields for details.

16. RecoveryDetails (O) – Information about when and where an instrument was recovered. Contains the same fields as DeploymentDetails.

17. Sensors (M) – A record describing the sensors on the instrument. These fields describe the sensors used to collect data, they do not contain the data themselves. Limited support for archiving low bit-rate sensor information is provided in the Data field (14).

   a. ReferencePoint (ᴀA C) – Specification of the reference point used for this instrument. Sensor geometry measurements are relative to this point. Example: “Geometry is measured from the center of the instrument.” When Geometry is specified in Sensor, Audio, or Depth, ReferencePoint is required.
   b. Sensor (O $+\infty$) – A generic sensor. Contains the following fields:
      1. Number ($\mathbb{Z}^+$ M) – Sensor index, must be unique if Sensor is repeated.
      2. SensorId (ᴀA M) – A value that uniquely identifies this sensor, such as a serial number.
      3. Geometry (O) – The geometry group of fields describe how the sensor is oriented relative to the defined ReferencePoint. Subfields:
         - x_m ($\mathbb{R}$ M) – Meters from the designated point along the X axis of the sensor’s transverse plane. When looking from the rear of the instrument to the front, positive values are to the right and negative values to the left.
         - y_m ($\mathbb{R}$ M) – Meters from designated point along the Y axis of the sensor’s transverse plane. Positive values are towards the front of the instrument and negative values towards the rear.
         - z_m ($\mathbb{R}$ M) – Meters from the designated point along the Z axis of the sensor’s sagittal plane. Positive values are above the center point and negative values below.
      4. Name (ᴀA O) – Sensor name, this could be a model number or user-friendly identifier.
      5. Description (ᴀA O) – Description of the sensor.
   c. Audio (O $+\infty$) – A sensor designed for recording audio. Contains generic sensor fields above and:
      1. TransducerId (ᴀA O) – Transducer identifier.
      2. PreampId (ᴀA O) – Preamplifier identifier
   d. Elevation (ᴀA O) – A depth/elevation sensor. Contains the same fields as a generic sensor.

18. MetadataInfo (O) – Information about maintenance of the fields contained in this record. See common elements section for details.

Detections

A detection record should be generated when an algorithm or analyst has examined the acoustic data for phenomena of interest. For systematic detections, we assume that all data between the start and end of the analysis effort period have been examined with the exception of aperiodic or periodic intervals documented within the record. Each detection record contains:

1. Id (ᴀA, M) – Text that uniquely identifies this set of detections.
2. Description (O) – A record containing optional prose about the objectives, abstract, and method associated with this set of detections. See Common Data Elements (item 2, p. 27) for further details.

3. DataSource (M) – A record with an EnsembleId **or** DeploymentId. Exactly one of the fields below is required to link the set of detections with information about the recording.
    a. DeploymentId (ᴀA M) – Matches an Id contained in a Deployment record. Detections records should not be accepted if the DeploymentId does not match a deployment record.
    b. EnsembleId (ᴀA M) – Matches an Id contained in an Ensemble record. Detections records should not be accepted if the EnsembleId does not match an ensemble record.

    Constraints: Exactly one of DeploymentId or EnsembleId must be present.
4. Algorithm (M) – A record that describes how the detections were obtained. The record is used in multiple schemata and is documented in the common data elements section (item 3, p. 27).
5. QualityAssurance (O) – A record describing the process used to ensure that the information contained within the detections record is correct. Contains:
    a. Description (O) – A record containing textual descriptions of Objectives, Abstract, and Method. See Common Data Elements, item (item 2, p. 27) for details.
    b. ResponsibleParty (O) – A record indicating who was responsible for the quality assurance and how to contact them. See Common Data Elements, item 5, p. 28.
6. UserId (ᴀA M) – An identifier for the user who submitted the set of detections.
7. Effort (M) – A record that describes for what the analyst or algorithm was searching. It contains the following fields:
    a. Start ($\overrightarrow{🕒}$ M) – List of analysis starts (date and time).
    b. End ($\overrightarrow{🕒}$ M) – List of analysis ends (date and time).

    Together, these describe periods during which analysis is in effect. Constraints: Start and End must be chronologically ordered and have the same length. Each Start must be before its corresponding End. Start and End periods may not overlap. Each Start and End must be between the Start and End of the deployment or ensemble referenced in DataSource.

8. dBReferenceIntensity_uPa ($\mathbb{R}$ O) – Provides a reference intensity in µPa for decibel measurements. This is nominally 1 µPa for underwater measurements and 20 µPa for in-air measurements.
9. Kind (M +∞) – One or more records containing descriptions of the signals that are being systematically detected/measured.
    a. SpeciesId (M $\mathbb{Z}$) – Positive numbers represent taxonomic ranks as per the integrated taxonomic information system (ITIS, https://itis.gov) which assigns taxonomic serial numbers to species. Anthropogenic signals are recorded as *Homo sapiens* (taxonomic serial number 180092). Abiotic phenomena are recorded with the reserved number of -10, sounds of unknown origin are denoted by the number -20.
    b. Call (ᴀA O) – Information about the type of sound recorded. For biotic sounds, this is a call type, e.g., for blue whales one might use "D" to denote the stereotyped down swept D call. As there is a lack of consensus within the bioacoustics community with respect to the naming of calls, we do not mandate specific names for calls although this would be certainly helpful. For a list of suggested names, see the supplemental information in Roch *et al.* (2016). We recommend anthropogenic sounds be labeled with call names representing the sound, such as "ship" or "sonar." Abiotic sounds should be labeled with a call representing the phenomena, such as "rain" or "earthquake."

c. Parameters (O) – Each detection may be characterized by a set of parameters or measurements. This record permits specification of a limited number of measurement fields that indicate specific measurement effort.
    - Subtype (ᴀA O) – A call subcategory, such as a known variant of a stereotyped call. Recording this field in effort enables systematic analysis of acoustic differences due to variations that may be related to social learning, population-level differences, etc.
    - FrequencyMeasurements_Hz ($\vec{\mathbb{R}}$ O) Provides a list of frequencies at which measurements are made. The FrequencyMeasurements_dB (item 18.5, p. 17) field records vectors of measurements at these frequencies. Measurement methodology should be detailed in Algorithm (item 4, p. 15).

d. Granularity (M) – Record that indicates what kind of detections are being made. Must be one of the following values:
    - "binned" – Presence/absence is reported within bins. When binned is selected, the following fields are valid:
        - BinSize_min ($\mathbb{R}$ M) - Duration in minutes of presence/absence bins
        - FirstBinStart (🕑 O) – Bins are assumed to start at the beginning of the specified effort. FirstBinStart permits an alternative timestamp to be used. This field allows users to make bin starts consistent. For example, if BinSize_min was 10 and the analysis started at 2022-03-15T15:17:00Z, one could set bins to align with hour and day boundaries: 2022-03-15T15:20:00Z.

        Constraints: FirstBinStart should not be later than BinSize_min after the start of analysis.

        Caveats: Analysis software might not take into account lost effort when FirstBinStart is after the start of detection effort.

    - "call" – each call is individually annotated in the detections
    - "encounter" – Detections report a set of calls, such as the period between the start of toothed whales starting to echolocate and when they stop or move out of detection range. When encounter is selected, the following field may be used:
        - EncounterGap_min ($\mathbb{R}$ O) – The minimum amount of time between two calls such that they are considered to be in separate encounters.
    - "grouped" – The detection represents a group of calls with a structural relationship. Examples include song and echolocation click trains.

10. OnEffort (M) – A record containing detections that were made systematically. All OnEffort detections must be represented in the specified Effort. The OnEffort record consists of a list of Detection (O +∞) records. Each Detection record contains:
    a. Input_File (ᴀA O) – Name or URI of file containing audio data associated with this detection.

b. Start (🕑 M) – Time and date of the beginning of the detection.
c. End (🕑 O) – Time and date the detection is complete.
d. Count ($\mathbb{Z}^+$ O) – For binned-, encounter-, and grouped-granularity detections, an optional number of calls associated with the detection, e.g., there were 8,957 observable echolocation clicks within the encounter. Times of individual calls are not recorded in this record.
e. Event (ᴀA O) – An event tag that can be used to refer to a detection.
   Constraint: Must be unique within the set of Detections.
f. UnitId ($\mathbb{Z}^+$ O) – Only used when the DataSource is an ensemble. Used to denote which deployment within the ensemble the detection is associated with.
g. Channel ($\mathbb{Z}^+$ O) – Specifies the channel of the recording that was used to produce the detection. When ensembles are used, refers to the channel within the specified ensemble unit (UnitId). If the detection is dependent on multiple channels (e.g., detection on a beamformed signal), Channel should be omitted.
h. SpeciesId ($\mathbb{Z}$ M) – Taxonomic serial number within the Integrated Taxonomic Information System (https://itis.gov) for biotic or anthropogenic signals, or -10 for abiotic signals. Anthropogenic signals are recorded with the taxonomic serial number for *Homo sapiens*.
i. Group (ᴀA O) – Species-level text for describing population-level groupings, such as distinguishing between the Shio- and Naisa-type short-finned pilot whales.
j. Call (ᴀA O +∞) – Call type associated with the detection, may be repeated if more than one type of call is detected within the Start/End interval (applicable to non-call detection granularities).
k. Parameters (O) – Structure documenting parameters/measurements of the detection. In many cases, these parameters can be measured in various ways, such as the computation of signal to noise ratio, received level, frequency measurements, etc. Users of the standard should populate the Algorithm record with enough details to make the measurements unambiguous. The Parameters record contains:
   1. Subtype (ᴀA O) – Call subcategory.
   2. Score ($\mathbb{R}$ O) – Measurement from a detection/classification algorithm related to the classification decision, e.g., a likelihood ratio or distance from a classification boundary.
   3. Confidence ($\mathbb{R}$ O) – A measurement of confidence that the classification/detection decision was correct. Higher numbers indicate greater confidence.
      Constraint: Must be in the interval between 0 and 1 inclusive.
   4. QualityAssurance (ᴀA O) – One of three text values:

      "unverified" – The detection has not been quality controlled,

      "valid" – The detection has been validated by the quality control process.

      "invalid" – The detection failed to be validated by the quality control process.

   5. ReceivedLevel_dB ($\mathbb{R}$ O) – Received level of the call in dB relative to the ReferenceIntensity_uPa declared in the Effort section.

6. FrequencyMeasurements_dB ($\vec{\mathbb{R}}$ O) – Vector of frequency measurements that correspond to the frequency list in FrequencyMeasurements_Hz for the SpeciesId (and Call if present) specified in the Effort record.
7. SNR_dB ($\mathbb{R}$ O) – Ratio of signal to noise (Signal – Noise as dB is a log unit).
8. MinFreq_Hz ($\mathbb{R}$ O) – Minimum frequency of signal.
9. MaxFreq_Hz ($\mathbb{R}$ O) – Maximum frequency of signal.
10. PeakFreq_Hz ($\mathbb{R}$ O) – Peak frequency of signal.
11. Peaks_Hz ($\vec{\mathbb{R}}$ O) – Vector of frequencies within a signal that comprise local maxima.
12. Duration_s ($\mathbb{R}$ O) – While not needed for detections of individual calls with a Start and End time, this measurement can be used to denote the mean duration of calls when detections cover multiple calls, such as the presence-absence detections of the "binned" Granularity effort.
13. Sideband_Hz ($\vec{\mathbb{R}}$ O) – Vector of signal sideband frequencies sorted from lowest to highest frequency.
14. Tonal (O) – A record of time and frequency measurements consisting of:
    - Offset_s ($\vec{\mathbb{R}}$ M) – Offsets from start time in seconds.
    - Hz ($\vec{\mathbb{R}}$ M) – List of frequencies for each offset.
    - dB ($\vec{\mathbb{R}}$ O) – Energy at each offset in dB.

    Constraints: All vectors present in the tonal measurement must be of the same length.
15. EventRef (ᴀA O +∞) – Reference to other event identifiers within this set of detections (10.e, p. 17). Can be used to specify detection structure, such as referring to the individual notes in a song phrase.
16. UserDefined (ᴀA O) – Arbitrary text to provide extensibility. It is suggested that some form of structure be used such as JSON, XML, etc.

l. Image (ᴀA O) – An identifier, such as a URL or DOI that lets a user retrieve an image of the detection, such as a spectrogram plot.
m. Audio (ᴀA O) – An identifier, such as a URL or DOI that lets a user retrieve a short segment of audio containing the call.
n. Comment (ᴀA O) – Analyst comments on the detection.

11. OffEffort (O) – A record containing non-systematic detections. The record contains a sequence of Detection records with the same content as in the OnEffort record. These records represent opportunistic detections that are unlikely to be characteristic of the results that one would observe with systematic effort.
12. MetadataInfo (O) – Structure containing information about this record. Contains:
    a. Contact (M) – Details on how to contact the person responsible for maintaining the record. See ResponsibleParty in common data elements (5, p. 28) for details about contact information.
    b. Date (🕑 M) – Timestamp of last update to the record.
    c. UpdateFrequency (ᴀA M) – Indication of planned update frequency. Must be one of the following values:
        - "as-needed"

- “unplanned”
- “yearly”

Localizations

Localization records capture spatial information about the direction, position, or motion of a sound source. A localization record should be generated when directional or positional analysis has been conducted over a section of a deployment. Each localization record contains:

1. Id (AA, M) – Text that uniquely identifies this set of location information.
2. Description (O) – A record containing optional prose about the objectives, abstract, and method associated with this set of detections. See Common Data Elements (item 2, p. 27) for further details.
3. DataSource (M) – A record with an EnsembleId **or** DeploymentId. Exactly one of the fields below is required to link the set of localizations with information about the recordings.
    a. DeploymentId (AA C) – Matches an Id contained in a Deployment record. Localization records should not be accepted if the DeploymentId does not match a deployment record.
    b. EnsembleId (AA C) – Matches an Id contained in an Ensemble record. Localization records should not be accepted if the EnsembleId does not match an ensemble record.

   Constraints: Exactly one of DeploymentId or EnsembleId must be present.
4. Algorithm (M) – A record that describes how the direction or position information was estimated. The Algorithm record contains a number of mandatory and optional fields that are used in multiple schemata. Algorithm is documented in the common data section (item 3, p. 27).
5. QualityAssurance (O) – A record describing the process used to ensure that the information contained within the set of localizations is correct. Contains:
    a. Description (O) – A record containing optional textual (AA O) descriptions of Objectives, Abstract, and Method for the quality assurance process. See Common Data Elements (item 2, p. 27) for details.
    b. ResponsibleParty (O) – A record indicating who was responsible for the quality assurance and how to contact them. See Common Data Elements (item 5, p. 28).

   Note that this QualityAssurance record only describes the process used for quality assurance, individual records containing position or direction information are marked individually with information about the results of the quality assurance process.
6. UserId (AA M) – An identifier for the user who submitted the set of localizations.
7. Effort (M) – A record that provides details of the periods for which analysis was conducted, the types of direction or location information being produced, and how the data should be georeferenced. It contains the following fields:
    a. Start (M) – List of analysis start dates and times.
    b. End (M) – List of analysis end date and times.

    Together, these describe periods during which analysis was conducted.

Constraints: Start and End must be chronologically ordered and have the same length.   Each Start must be before its corresponding End. Start and End periods may not overlap. Each Start and End must be between the Start and End of the deployment or ensemble referenced in DataSource.

c. CoordinateReferenceSystem (M) – A record that defines how location or direction information is recorded within the current set of localizations. This specification conforms to a subset of ISO 19111:2019. The type of coordinate system is specified by Subtype and Name.
    1. Subtype (ᴀA M) – Must be one of the following values:
        - Geographic – A coordinate reference system with an ellipsoidal projection system.
        - Derived – A coordinate system that is derived from another coordinate reference system.
        - Engineering – A coordinate system that provides an anchor reference relative to a local position and does not consider the curvature of the Earth.
    2. Name (ᴀA M) –Specifies a specific coordinate reference system for a given family of reference systems specified in Subtype. Valid names:

        Subtype: Geographic

        - WGS84 – Data will conform to the WGS 84 standard, which specifies positions with longitude and latitude.

        Subtype: Derived:
        Derived coordinate systems are relative to a ReferenceFrame (item 7.c.3, p. 21).

        - UTM – Data will conform to the Universal Transverse Mercator measurement system that provides a local flat-Earth projection relative to a WGS 84 reference point (Snyder, 1987). UTM uses a set of false Eastings and Northings in meters that are translated such that they are always positive numbers with respect to the local reference point. Each projection is referred to as a zone.

        Subtype: Engineering:
        All engineering measurements are angles measured in degrees and/or distances measured in meters. Measurements are relative to a datum, a reference point/surface, specified by the ReferenceFrame (item 7.c.3, p. 21). The following Name values are valid for the Engineering Subtype:

        - Cartesian – Two or three-dimensional points (m) relative to a datum.
        - Polar – Polar coordinate systems consist of a planar angle measured in degrees counterclockwise to a reference axis within a plane, and potentially a radial distance d (m).

- Spherical – Measurements consist of a pair of angles measured in counterclockwise degrees. $\angle_1$ measures the angle between the planar reference axis and the source, and $\angle_2$ is the angle relative to the plane at the datum with positive values indicating that the localized sound producer is above the plane. An optional distance d (m) along the resulting direction may also be specified.
- Cylindrical – Coordinates will consist of triplets of values consisting of the angle measured in counterclockwise degrees relative to the planar reference axis, a distance along the radial $d_\angle$ (m), and a vertical distance relative to the reference plane $d_\perp$ (m). Positive distances imply that the sound source is above the plane.
- Range – Straight line distance (m) with no orientation information recorded.
- PerpendicularRange – Perpendicular distance to an effort track line (m). Effort tracks are not infinite, and if the intersection between the sound source and the effort segment does not exist, a perpendicular range should not be recorded.

3. ReferenceFrame (C) – Describes the datum, required for Subtype values Derived and Engineering; **omitted for Subtype Geographic** which has an implicit datum.
    1. Anchor (ᴀA M) – Describes the type of datum used. Valid values:
        - WGS84 – Reference is specified in WGS 84
        - Instrument – Reference is relative to instrument. Valid only for Subtype=Engineering.
        - UTMZone – Reference is specified as a UTM zone indicator (which encodes a WGS 84 position). Valid only for Name=UTM.

        When the Anchor is WGS84 or UTMZone, elevation references are relative to average sea level.

    2. Latitude ($\mathbb{R}$ C) – WGS 84 latitude, required when Anchor=WGS84
    3. Longitude ($\mathbb{R}$ C) – WGS 84 longitude, required when Anchor=WGS84
    4. UTMZone (ᴀA C) – Text that must match a valid UTM zone indicator. UTM zones consist of a number between 1 and 60 followed by a northern (N) or southern (S) hemisphere reference, e.g. 17S. Required when Anchor=UTMZone.
    5. Datum (ᴀA C) – Textual description of the datum of the coordinate system. See the scenarios appendix for examples of this. Datum is required when SubType=Engineering and Anchor=Instrument. Other Anchor types specify their datum using Latitude/Longitude or UTMZone.
4. LocalizationType (ᴀA M+∞) – Indicates what type of localization effort was made and must be one or more of the following values:
    - Bearing – Localizations reflect a direction towards a sound source relative to the datum, but not the distance. Planar angular measurements are made relative to the longitudinal axis of instrument when the

ReferenceFrame's Anchor is Instrument. For Anchors WGS84 and UTM, planar angles are measured relative to the true north vector passing through the anchor location.

- Range – Straight line distance (m) from instrument reference point to the sound source with no orientation information. Used with Engineering / Range.
- PerpendicularRange – Perpendicular distance to a track line associated with the instrument's motion vector which is assumed to be a straight line. Used with Engineering / PerpendicularRange.
- 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 associated with a moving sound source.

While LocaliztionType can be repeated, values are subject to the following restrictions:

- No type may be repeated.
- With the exception of Range and PerpendicularRange, all localization entries must be in the specified CoordinateReferenceSystem. Range and PerpendicularRange are exceptions as they are single dimension units (m), and it may be convenient to store these values as they are usually known at the time the location information is estimated.

5. TimeReference (ᴀA M) – Localizations contain timestamps, this field indicates how they should be interpreted:
   - absolute – Timestamps represent the time at which the sound was produced (e.g., derived from the estimated location, sound propagation model, and time of arrival).
   - channel – All times are presented relative to arrival on specified channels. Individual localizations will contain information that indicates the reference channel, see TimeReferenceChannel and TimeReferenceEnsembleUnit below.
   - relative – Times may be relative to any channel of opportunity within the recording. For most applications, this is the recommended value.
6. Dimension (ℤ M) – Each localization may vary in the dimensionality of the coordinate space, e.g., if the signal is not detected on enough transducers, we may only be able to produce a two-dimensional localization instead of a three-dimensional one. Dimension indicates the **maximal number** of dimensions that will be used. For example, LocalizationType=Point may have a Dimension of 2 or 3, Bearing would be 1 or 2, Range would be 1.

8. Localizations (M) – A collection of Localization records describing directional or positional information. Each Localization contains the following fields:

a. Event (ᴀA O) – An event identifier. When present, it consists of text that must be unique within the set of localizations associated with the current localization Id. The Event and Id together can be used to reference a specific localization.
b. TimeStamp (🕑 M) – The date and time associated with the positional or directional measurement. The time refers to either the time the sound was generated, an approximation of the time the sound arrived at the receiver, or the arrival time on a specific transducer as specified by the TimeReference field in the Effort record.
c. SpeciesId (ℤ O) – Most localizations should refer back to the detection record from which the species producing the sound can be obtained. In cases where users wish to record localizations without recording the detections, recording the species (if available) can be helpful, and this field allows the recording of an ITIS taxonomic serial number if so desired.
d. References (O) – A record that contains optional fields describing the timestamp and detections/localizations used to create this localization record.
    1. TimeReferenceChannel – (ℤ C) – When the TimeStamp (8.b) refers to a specific channel (TimeReference=channel), this field contains the channel number.
    2. TimeReferenceEnsembleUnit (ℤ C) – When the TimeStamp (8.b) refers to a specific channel (TimeReference=channel) and the DataSource is an ensemble of deployments, the ensemble unit number indicates to which deployment the TimeReferenceChannel refers.
    3. Reference ($\overrightarrow{\text{ᴀA}^3}$ C) – List of detections or localizations that contributed to this localization record. Example: A list of bearings that were crossed to produce a two-dimensional location. Each reference must specify (type, id, event) where type is “localizations” or “detections”, Id identifies a record of the specified type, and event indicates a specific localization or detection as matched on the Event field of the indicated record..

Subtype	Name	LocalizationType	Valid Localization Records
Geographic	WGS84	Point	WGS84Coordinate
		Track	Track of WGS84CoordinateList
Derived	UTM	Point	UTMCoordinate
		Track	Track of UTMCoordinateList
Engineering	Cartesian	Point	CartesianCoordinate
		Track	Track of CartesianCoordinateList
	Polar	Bearing	Bearing
		Point	AngularCoordinate
		Track	Track of AngularCoordinateList
	Spherical	Bearing	Bearing
		Point	AngularCoordinate
		Track	Track of

			AngularCoordinateList
	Cylindrical	Point	CylindricalCoordinate
		Track	Track of CylindricalCoordinateList
Any	Any	Range	Range
		PerpendicularRange	PerpendicularRange

One of the fields below will be used to record location information. In most cases, an optional error field allows specifying the standard error for the measurement.

e. Coordinate ($\mathbb{R}$, $\mathbb{R}^2$ or $\mathbb{R}^3$ O) – A one-to-three-dimensional coordinate (x), (x, y), or (x, y, z). Interpretation of Coordinate depends on the coordinate reference system:
    - Geographic / WGS84 : (x=longitude (°), y=latitude (°), z=elevation (m))
    - Derived / UTM: (x=false easting (m), y=false northing (m), z=elevation (m))
    - Engineering / Cartesian: (x=offset (m), y=offset (m), z=offset (m)). Interpretation of axes is established by the ReferenceFrame declared in the Effort's CoordinateReferenceSystem (7.c, p. 20).
    - Engineering / Range: (x=offset (m)). Distance to sound source.
    - Engineering / PerpendicularRange (x=offset (m)). Distance to sound source from a line perpendicular to the instrument's track line.

    Each Coordinate may be followed by a CoordinateError ($\mathbb{R}$, $\mathbb{R}^2$ or $\mathbb{R}^3$ O) that specifies the standard error of each coordinate.
f. Bearing ($\mathbb{R}$ or $\mathbb{R}^2$ O) – A one- or two-dimensional bearing, $(\angle_1)$ or $(\angle_1, \angle_2)$. , $\angle_1$ is measured in counterclockwise degrees (°) relative to the true north when $\text{Anchor} \in \{\text{WGS84}, \text{UTM}\}$ or the instrument's longitudinal axis when Anchor=Instrument.
    Each Bearing may be followed by a BearingError ($\mathbb{R}$ or $\mathbb{R}^2$ O) that specifies the standard error of each angle.
g. AngularCoordinate ($\mathbb{R}^2$ or $\mathbb{R}^3$ O) – A 2D location based on a planar angle (°) and distance (m) $(\angle_1, d)$ or a 3D location from a planar angle (°) , azimuth (°) and distance (m): $(\angle_1, \angle_2, d)$. Angles are measured in the same manner as for the Bearing record.
    Each AngularCoordinate may be followed by an AngularCoordinateError that specifies the standard error of each angle and distance.
h. CylindricalCoordinate ($\mathbb{R}^2$ or $\mathbb{R}^3$ O) – A 2D/3D coordinate based on a planar angle (°), distance $d_\angle$ along that angle (m), and distance perpendicular $d_\perp$ to the plane (m): $(\angle_1, d_\angle)$ or $(\angle_1, d_\angle, d_\perp)$.
    Each CylindricalCoordiante may be followed by a CylindricalCoordinateError ($\mathbb{R}^2$ or $\mathbb{R}^3$ O) that specifies the standard error of the angle and the distances.
i. Track – A record that specifies a collection of localizations belonging to the same acoustic source. Information within a track depends on the type of coordinate reference system that is being used.

    For coordinate reference systems that record Coordinate records, three fields are used:

1. CoordinateList ($\overrightarrow{\mathbb{R}^2/\mathbb{R}^3 \text{🕑}}$ M) – A list of temporally ordered and timestamped coordinate points.
2. CoordianteListError ($\overrightarrow{\mathbb{R}^2/\mathbb{R}^3}$ O) – Optional list of standard errors corresponding to each entry of the CoordinateList.
3. CoordinateBounds (M) – Spatial bounding box specifying the northwest and southeast extents of the measurements. For Euclidean engineering coordinates, northwest should be interpreted as min(x), max(y), and southeast max(x), min(y). Elevation bounds show the min and max elevation. When a bounding box crosses the WGS 84 Meridian, it is recommended to specify unwrapped degree values. Subfields:
    1. NorthWest ($\mathbb{R}^2$ M) – Most NorthWest Coordinate in Track.
    2. SouthEast ($\mathbb{R}^2$ M) – Most SouthEast Coordinate in Track.
    3. Elevations ($\mathbb{R}^2$ O) – Lowest and highest elevation in Track.

Angular and cylindrical coordinates have similar fields although no CoordinateBounds are provided as CoordinateBounds are most useful when Tracks are recorded in WGS 84.

Angular coordinate reference systems use:

4. AngularCoordinateList ($\overrightarrow{\mathbb{R}^2/\mathbb{R}^3 \text{🕑}}$ M) – A list of temporally ordered angle(s) and distances, each accompanied by a timestamp: $(\angle_1, d, \text{🕑})$ or $(\angle_1, \angle_2, d, \text{🕑})$.
5. AngularCoordinateListError ($\overrightarrow{\mathbb{R}^2/\mathbb{R}^3}$ O) – Optional list of standard errors corresponding to each entry of the AngularCoordinateList.

Cylindrical coordinate reference systems use:

6. CylindricalCoordinateList ($\overrightarrow{\mathbb{R}^3 \text{🕑}}$ M) – where each temporally ordered entry consists of $(\angle_1, d_\angle, d_\perp, \text{🕑})$
7. CylindricalCoordinateListErrror ($\overrightarrow{\mathbb{R}^3}$ O) – Optional list of standard errors corresponding to each entry of the CylindricalCoordinateList.

j. InstrumentTelemetry (O) – Deployment records capture some instrument telemetry information, but these are frequently downsampled to avoid excessive storage requirements. The optional InstrumentTelemetry field permits recording high resolution readings from instruments without having to record a complete high-resolution copy of the telemetry. It consists of the following optional fields:
    1. Position (🌎 O) – Position in a standard coordinate system.
    2. Bearing_DegN ($\mathbb{R}$AA O) – Bearing of instrument in degrees relative to north and a text field that indicates if the north reference was based on “true” or “magnetic” north.
    3. Speed ($\mathbb{R}$AA O) – Speed of instrument and text specifying units. Must be knots, m/s, or km/s.
    4. Pitch_deg ($\mathbb{R}$ O) – Instrument pitch in degrees [-90, 90] relative to the lateral plane. Positive numbers indicate that the bow (front) is oriented above the lateral plane.

    5. Roll_deg (ℝ O) – Instrument roll in degrees (-180, 180] relative to the longitudinal plane. Positive numbers indicate that starboard (right relative to front) side of the instrument has tilted up.
    6. Elevation_m (ℝ O) – Elevation (m) of animal or other localized source above or below sea level.
    7. GroundElevation_m (ℝ O) – Elevation (m) of the ground or sea/lake/river bed. Elevations in water are usually negative except in the case of bodies of water above sea level.
  k. QualityAssurance (ᴀA O) – Indicates if the localization have been verified. Must be one of: unverified, valid, or invalid.
  l. Parameters –WHAT MIGHT BE COMMON ACROSS APPLICATIONS + UserDefined?
9. BespokeData (O) – Record with custom data supporting the set of localizations (e.g., data related to the method such as cross correlations, etc.) See Common Data Elements item 7, p. 29, for details.
10. MetadataInfo (O) – Record describing the person responsible for maintaining this record, the date the record was last modified, and the plan for future updates. See Common Data Elements item 6, p. 29.

Calibrations

Calibration records are designed to capture information about acoustic sensors. The standard supports multiple calibrations of the same instrument as well as calibrations entire instruments or components such as transducers (hydrophones/microphones) and preamplifiers. Specification of methods for calibration are beyond the scope of this standard.

Each calibration record shall have the following information:

1. Id (ᴀA M) – The Id field contains text that identifies a specific instrument, preamplifier, or transducer. Unlike other fields in these schemata, the Id field need not be unique as instruments can be calibrated multiple times.
2. TimeStamp (🕑 M) – The date and time at which the calibration was performed.
3. Type (ᴀA M) – One of the following text values that indicates what has been calibrated:
    - transducer – Calibration of a microphone or hydrophone.
    - preamplifier – Calibration of a preamplifier circuit.
    - recorder – Calibration of a recording unit independent of any external connections (e.g., transducer).
    - end-to-end – Calibration of an assembled instrument.
4. Process (O) – A record that contains the same fields as the Algorithm record specified in the common data elements section (see p. 27).
5. ResponsibleParty (O) – Information on the person or organization responsible for the calibration. Conforms to the ResponsibleParty element in the common data elements section (p. 28).
6. QualityAssurance (O) – A record describing whether or not the calibration has been validated. When present, it contains the following fields:
    a. Quality (ᴀA M) – Text with one of the following values:
        - valid – Calibration has been validated
        - invalid – Calibration is known to have errors

        - unverified – Calibration has been performed, but has not been validated.
    b. Comment (ᴀA O) – Text with any descriptive information about the quality assurance.
    c. AlternateCalibration (ᴀA O) – A record which describes a calibration that can be used as a proxy for this one when a calibration is faulty. Contains two fields:
        a. Id (ᴀA M) – A text string that identifies one or more calibration records by their Id. In general, it is expected that this is referring to a calibration of the same Type, although this may not always be available.
        b. TimeStamp (🕑 O) – A timestamp that may be used to refer to a specific calibration date for the proxy.
7. IntensityReference_uPa ($\mathbb{R}$ O) – Reference intensity for decibel measurements, typically 1 for in-water measurements and 20 for in-air measurements.
8. SensitivityReference (ᴀA O) –Text specifying a reference measurement for transducer sensitivity.
9. Sensitivity_dBV ($\mathbb{R}$ O) – Measurement of transducer sensitivity in dB/V across the region of flat response.
10. Sensitivity_VFS ($\mathbb{R}$ O) – Voltage required from source to produce full scale.
11. BitDepth ($\mathbb{R}$ O) - For analog to digital conversion, how many bits were…
12. FrequencyResponse (O) – Record indicating sensitivity at different frequency probes. Contains the following fields:
    a. Hz ($\vec{\mathbb{R}}$ M) – List of probe frequencies whose sensitivity was measured.
    b. dB ($\vec{\mathbb{R}}$ M) – List of sensitivities corresponding to the probe frequencies.
    c. ResponseType (ᴀA M) – One of the following values:
        - absolute – Frequency response dB values are absolute sensitivity
        - relative – Frequency response dB values are offsets to the sensitivity measurement.

Common Data Elements

1. Id (ᴀA)– The Id field contains a text string that serves as a key for any given deployment, ensemble, set of detections or localizations. Within a given category such as detections, it must be unique.
2. Description – A record containing the following text fields
    a. Objectives (ᴀA O) – Describes the objectives of the effort associated with the record.
    b. Abstract (ᴀA O) – A brief summary describing what was done.
    c. Method (ᴀA O) – A high-level description of the methods that were used.
3. Algorithm (M) – A record that describes how the data were obtained
    a. Method (ᴀA M) – Text-based description of the algorithm. Citations of published manuscripts are appropriate when feasible.
    b. Software (ᴀA M) – Name of software that supports the analysis. When human analysts are conducting analysis, this is frequently the software that they are using to aid in conducting the analysis, e.g., Raven Pro (K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology). Many systems for automated analysis support a plug-in architecture that allows additional algorithms to be used. In these

cases, the algorithm responsible for analysis, the “plug-in,” should be documented, with the underlying system documented in the SupportSoftware section.

   c. Version (ᴀA O) – Software version identifier. Software that is in development does not always have a version number assigned. When version control systems are being used (e.g., git, mercurial, cvs, etc.) one can use the identifier associated with the commit of the version that was used. While this field is optional, it is strongly encouraged, or results will be difficult to reproduce.
   d. Parameters (ᴀA M) – The set of parameters used during analysis. This field is unstructured as parameters vary between algorithms. We strongly recommend to use a consistent way of describing parameters that captures all settings and is searchable by algorithms. We recommend embedding parameters in the structured text provided by extensible markup language (XML) or JavaScript object notation (JSON) as these are easily parsable by software and hence searchable.
   e. SupportSoftware (O +∞) – A record that may be repeated multiple times to describe any additional software that was required for the analysis. Each record contains:
      i. Software (ᴀA M) – The name of the software, e.g., PAMGuard
      ii. Version (ᴀA O) – Software version identifier, e.g. 2.0.28
      iii. Parameters (ᴀA M) - Any parameters that were used with the support software.

   Allowing SupportSoftware to be repeated permits users to denote multiple dependencies, such as noting that they used Java version 8u20 with PAMGuard on Windows 11. It is at the user’s discretion as to what level of support software should be documented.

4. Person – Based on a subset of the OpenGIS Sensor ML standard, describes an individual.
   a. id (ᴀA O) – A handle to an external database. This is useful for exposing data publicly in settings where privacy concerns require the obfuscation of individual’s names. In such cases, implementations may choose not to expose sensitive data and return a key derived from the database and person entry.
   b. surname (ᴀA M) – Last or family name
   c. name (ᴀA M) – First name
   d. userID (ᴀA M) – Account name associated with person.
   e. affiliation (ᴀA M) – Organization with which the person is affiliated.
   f. phoneNumber (ᴀA M) – Textual representation of a telephone number, e.g. +1 555-589-2983.
   g. email (ᴀA M) – Electronic mail address of individual.
5. ResponsibleParty – Based on the ISO 19115 specification, this contains the following:
   a. id (ᴀA O) – A handle to an external database. This is useful for exposing data publicly in settings where privacy concerns prevent the release of individual’s names.
   b. individualName (ᴀA O) – Person’s name.
   c. organizationName (ᴀA O) – Company, agency, or other organizational name.
   d. positionName (ᴀA O) –Role or title of the responsible party.
   e. contactInfo (O) – A set of contact information. Contains:
      i. phone (O) – Contains a set of contact numbers
         1. voice (ᴀA O +∞) – Textual representation of a telephone number, e.g. +1 555-589-2983.

            2. facsimile (ᴀA O +∞) – Textual representation of a facsimile number.
        ii. address (O) – Set of elements describing addresses for delivery or electronic correspondence.
            1. deliveryPoint (ᴀA +∞ O) – Company location code, “Attention of” message, other additional delivery instructions.
            2. city (ᴀA O) – Delivery city
            3. administrativeArea (ᴀA O) – Delivery area, such as state, department, or province.
            4. postalCode (ᴀA O) –Postal service delivery aid, e.g., United State zip code or French postal code.
            5. country (ᴀA O) – Delivery country.
            6. electronicMailAddress (ᴀA O) – Address for electronic delivery.
        iii. hoursOfService (ᴀA O) – Information on when responsibleParty can be reached.
        iv. contactInstructions (ᴀA O) – Information about contacting the person or organization detailed in the fields above.

6. MetadataInfo (O) – Information about maintenance of the fields contained in this record.
    a. Contact (M) – Contains information about the person and/or organization responsible for maintenance. See the description of ResponsibleParty in the Common Data Elements section (item 5, p. 28).
    b. Date (🕑 M) – Date and time of last update to this record.
    c. UpdateFrequency (ᴀA M) – Indicates frequency with which the record will be updated. Accepts the following values:
        i. as-needed – Updated as maintainer sees warranted.
        ii. unplanned – No current plans to update.
        iii. yearly – Updated on an annual basis.
7. BespokeData (O) – Custom data that is stored along with the record. Contains the following fields:
    a. Abstract (ᴀA M) – Description of the bespoke data.
    b. Data (O +∞) – Record with the following fields:
        i. URI (ᴀA M) - Either a uniform resource indicator indicating where the information can be found (e.g., web site, digital object identifier, etc.) or a filename. If a filename is specified, the file should be stored and managed by the data information system responsible for managing the records. The data information system must provide methods to retrieve such data, but interpretation of the data is the user’s responsibility.
        ii. Comment (ᴀA O) – Text describing this specific data resource if it is not obvious from the Abstract.
    c. UserDefined (O) – A list of names and textual values (ᴀA, ᴀA) describing bespoke data that is simple enough to store in text fields.

References

Arms, C. R., Felischauer, C., Murray, K., Nappier, M., and Holdzkom, L. (2023). "Sustainability of Digital Formats: Planning for Library of Congress Collections," Accessed March 26, 2023. https://www.loc.gov/preservation/digital/formats.

Au, W. L., and Hastings, M. C. (2008). *Principles of Marine Bioacoustics* (Springer, New York), pp. 679

Berners-Lee, T. (2005). "Uniform Resource Identifier (URI): Generic Syntax," Accessed March 23, 2023. https://www.ietf.org/rfc/rfc3986.txt.

Bradbury, J. W., and Vehrencamp, S. L. (1998). *Principles of Animal Communication* (Sinauer Associates, New York)

Darwin Core Maintenance Group (2021). "List of Darwin Core terms. Biodiversity Information Standards (TDWG)," Accessed June 14, 2023. http://rs.tdwg.org/dwc/doc/list/2021-07-15.

Fujioka, E., Soldevilla, M. S., Read, A. J., and Halpin, P. N. (2014). "Integration of passive acoustic monitoring data into OBIS-SEAMAP, a global biogeographic database, to advance spatially-explicit ecological assessments," Ecol. Inform. 21. 59-73, doi:10.1016/j.ecoinf.2013.12.004.

Guan, S., Moustahfid, H., Milan, A., and Mize, J. (2014). "A Metadata Convention for Passive Acoustic Recordings, Version 1.0," US Integrated Ocean Observing System, Silver Spring, MD, pp. 47.

GWG World Geodetic System and Geomatics Focus Group (2014). "Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems," in *NSG Standards Registry* NGA.STND.0036_1.0.0_WGS84, National Center for Geospatial Intelligence Standards (NCGIS), National Geospatial-Intelligence Agency (NGA), Springfield, VA, USA, pp. 208.

Halpin, P. N., Read, A. J., Fujioka, E., Best, B. D., Donnelly, B., Hazen, L. J., Kot, C., Urian, K., LaBrecque, E., Dimatteo, A., Cleary, J., Good, C., Crowder, L. B., and Hyrenbach, K. D. (2009). "OBIS-SEAMAP The World Data Center for Marine Mammal, Sea Bird, and Sea Turtle Distributions," Oceanography 22(2). 104-115.

Hildebrand, J. A. (2009). "Anthropogenic and natural sources of ambient noise in the ocean," Mar. Ecol. Prog.-Ser. 395. 5-20, doi:10.3354/meps08353.

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

IOOS Biology (2018). "Framework for Defining the IOOS Contribution to Biological Observing," International Ocean Observing System Association, pp. 9.

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

Roch, M. A., Batchelor, H., Baumann-Pickering, S., Berchock, C. L., Cholewiak, D., Fujioka, E., Garland , E. C., Herbert, S., Hildebrand, J. A., Oleson, E. M., Van Parijs, S. M., Risch, D., and Širović , A. (2016). "Management of acoustic metadata for bioacoustics," Ecol Info 31. 122-136, doi:10.1016/j.ecoinf.2015.12.002.

Roch, M. A., Baumann-Pickering, S., Batchelor, H., Hwang, D., Sirovic, A., Hildebrand, J. A., Berchok, C. L., Cholewiak, D., Munger, L. M., Oleson, E. M., Van Parijs, S., Risch, D., and Soldevilla, M. S. (2013). "Tethys: a workbench and database for passive acoustic metadata," Oceans 2013, pp. 5 pp., doi:10.23919/OCEANS.2013.6741361.

Shannon, G., McKenna, M. F., Angeloni, L. M., Crooks, K. R., Fristrup, K. M., Brown, E., Warner, K. A., Nelson, M. D., White, C., Briggs, J., McFarland, S., and Wittemyer, G. (2016). "A synthesis of two decades of research documenting the effects of noise on wildlife," Biological Reviews 91(4). 982-1005, doi:10.1111/brv.12207.

Snyder, J. P. (1987). "Map Projections: A Working Manual," U.S. Geological Survey, Washington, DC, pp. 383, doi:10.3133/pp1395.

Stowell, D. (2022). "Computational bioacoustics with deep learning: a review and roadmap," PeerJ 10, pp. e13152, doi:10.7717/peerj.13152.

Walmsley, P. (2002). *Definitive XML Schema* (Prentice Hall PTR, Upper Saddle River, NJ), pp. 528

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Doring, M., Giovanni, R., Robertson, T., and Vieglais, D. (2012). "Darwin Core: an evolving community-developed biodiversity data standard," PLoS One 7(1), pp. 8, doi:10.1371/journal.pone.0029715.

Annex A: Data types

The standard does not prescribe the machine representation for most data. Data types shall be transformable to the standard formats described below to ease data interchange between systems.

Text data: Character text shall be translatable to the Unicode encoding system. Unicode is a standard for representing character data that is capable of representing all human languages; standards may be found at: https://www.unicode.org/. We recommend that text data be stored as Unicode when feasible.

Two types of numbers are defined. The first is integers, which should be capable of representing at least the range of integers between $[-2^{63}, 2^{63}-1]$, which is the range of approximately $\pm 9 \times 10^{19}$. This corresponds to a 2s complement representation of a 64-bit integer. Non decimal numbers shall be representable with IEEE standard for double-precision floating-point arithmetic, IEEE 754-2019. These formats are commonly supported by nearly all 32- and 64-bit computers, operating systems, and programming languages.

Timestamps may be stored in any format, but must be representable in universal coordinated time (UTC) and capable of being display according to the ISO-8601 date and time standard. ISO-8601 uses the Gregorian calendar and a 24-hour clock. For simplicity, systems should be capable of displaying the all-numeric representation of timestamps which orders time units from largest (year) to smallest (seconds). Dates are specified by the complete year, two-digit month, and two-digit day separated by hyphens. This is followed by a capital letter T to indicate time, and a time specification. The time specification uses a 24-hour clock with two digits each for hours, minutes, and seconds, separated by colons. The seconds may contain a period followed by fractional seconds. The internal time representation should be capable of representing time to at least the microsecond resolution. An optional Z denotes UTC, which will be assumed if not present. Systems that parse ISO-8601 dates when adding new data should be capable of reading timezone offset notation which consists of a plus or minus followed a specification of hours and potentially minutes that the time differs from UTC: ±[hh], ±[hh]:[mm], or ±[hh][mm]. A timezone offset of +00 shall be interpreted as UTC time, timezone offsets of -00 are not permitted. Two examples of ISO 8601 timestamps:

- `2010-08-27T23:58:03.3975Z` represents 397,500 µs past 11:58:03 PM UTC on August 27th, 2010.
- `2023-03-15T10:54:00-07:00` and `2023-03-15T10:54:00-07` both represent March 15th, 2023 at 10:54 AM Pacific daylight time (PDT). PDT is 7 hours behind UTC on this date. Once parsed, the UTC time would be stored and querying either timestamp would return: `2023-03-15T17:54:00Z`.

Absolute positional information should be encoded using a common geographic standard. For longitude and latitudes, we recommend using the world geodetic system (WGS 84) standard (GWG World Geodetic System and Geomatics Focus Group, 2014), with latitudes and longitudes represented by the IEEE standard for double-precision floating-point arithmetic, IEEE 754-2019. A common alternative is the universal transverse Mercator (UTM) coordinate system (Snyder, 1987), which divides the earth into a series of grids over which the curvature of the earth is negligible. Grid regions are identified as polar regions or patches that are identified by longitude and latitude spans referred to as zones denoted as a letter (ᴀA) and numeric bands ($\mathbb{Z}^+$) respectively. Non-polar regions are 6° of longitude with latitude bands generally covering 8° (with the notable exception of the northernmost band which is slightly wider). Positions within a grid are represented as Eastings and Northings in meters ($\mathbb{R}$). Eastings are always relative to the longitude midpoint of a band and are offset by 5×10^5 m, resulting in positive measurements. Northings are dependent on whether the measurement is in the northern or southern hemisphere. For northern hemisphere grids, Northings are relative to the equator. Southern hemisphere grids measure the Northing by subtracting the distance to the equator from 1×10^7 meters.