NOTES: Metadata Standardization Working Group Meeting. December 12th 2015.

**Animal Telemetry Data & Metadata Meeting**

**Date Tuesday 15th December 2015**

**ADT 1030-1430 AKDT 0530-0930, GMT 1430-1830, PDT 0630-1030**

**Location: Franciscan C Imperial Ballroom, Hilton, Union Square, San Francisco**

**Attendees:**

**In the room:** Trevor McIntyre, Nico DeBruyn (RSA), Dan Costa (UCSC/ATN), Philip Goldstein (USGS/OBIS), Barbara Block (STANFORD/ATN), Melinda Holland (WC), Ed Bryant (Fastlok), Josh London, Rob Harcourt, Mike Weise (ONR/ATN), Dana Belden, Frederic Ballieul (SARDI), Donna Kehoe(LOTEK), Mike Fedak, Lars Boehme, Phil Lovell (SMRU), Sam Simmons (MMC/ATN).

**On the Phone/Webinar:** Lenore Bajona, Fred Whoriskey, Jonathan Pye, Marta Mihoff (OTN), Mark Meinhoff, Stephanie Smedbol (VEMCO), Peter Walsh, Xavier Hoenner (IMAS/emII).

**Apologies:** Hassan Moustahfid (NOAA/ATN), Roger Proctor (IMAS/emII), Erik Hoy (Thelmabiotel), Baldur Sigurgeirsson (Star Oddi) Sean Hayes (NOAA/ATN)

**To be added to mailing list:** Bernie McConnell (SMRU)

**Goals for the meeting:**

1. **Review template on Animal Telemetry data/metadata**
2. **Develop plan to implement an acceptable convention data & metadata and services standards and best practices for animal satellite/archival telemetry**

**Actions- for all:**

1. All agreed to target the Ocean Sciences meeting in February as the next date to convene and further discussions from today. RH to organize venue and contact details
2. With regard to the metadata template (or exchange format or database schema as might be more appropriate descriptors). Agreed that everyone should get any additional comments in ASAP following the meeting to Xavier and the Hobart team so that they can try to revise the document according to comments received and circulate it again before the Ocean Sciences meeting. Agreed that underlying these comments is the intent to outline bigger picture while focusing on data sharing for now
3. Formation of Animal Borne Task team

**Meeting Overview and individual action items**:

Josh London started with a verbal summary of some of his comments on the metadata standard that had been shared before the meeting. He asked if we were really going for a metadata standard with this document as it seemed more like a database schema to him and Sam Simmons offered that perhaps an “**exchange format**” is a better way to describe it/use for reference.

Other comments and other main points that came up in subsequent discussions were:

* From Josh London Current template is missing a field/fields for ‘recovered data’ – either via satellite or for archival tags, particularly the diving behavior. (Xavier Hoenner: this template is based on AATAMS data).
* How will the firmware and software version information be stored as part of the metadata as well as how the tag was programmed?
* **Possible action**: Several parts of the table focus on animal information, suggested we focus on tag specific information at least initially b/c many folks have their own standards or established practices there already.
* The importance of recording deployment start and end dates and locations or each as applicable in the metadata was noted.
* Suggested we use existing standards where available/applicable.
* Suggestion to vendors that some sort of common format be developed to get data from the tags so that it is them more easily digestible into a common database(s).
* Dan Costa – the standard/what we are developing could serve both purposes – a ‘prescriptive database’ (or schema) for anyone that doesn’t have one already and also as an exchange format ensuring exchanged fields/formats are clearly identified.
* **ACTION** Xavier Hoenner: The focus of the exchange format/schema should be the tagging /deployment event and then the recovery event and there may be biological data (mass, age, sex, etc.) associated with both of these.
* The importance of tracking individual tags was also discussed. For example one animal or deployment/recovery event may involve several instruments on the same animal. The operational utility needs to be clear and unequivocal in the Metadata.
* Note from Sam Simmons: in Dan Costa’s lab they used to have a simple tag database that followed the deployment, recovery, rebatterying, etc. “events” in the life of a tag. **ACTION**- person from Costa lab to communicate the schema/database with Xavier Hoenner.
* For consideration: add information to the deployment event for a tag that records what the instrument set up was (sampling frequency, etc.)
* Lars Boehme raised the use of netCDF format used by oceanographers as of potential interest to this group to review with regard to the exchange format.
* Question was raised: in the exchange format/schema where is the line between metadata and “data” and don’t we want both in the exchange format/file?
* **ACTION**– Josh London trying to find a constructive place to “start” – Does it make sense to start with archival time series data and work from there? Is there a netCDF format for time series that we could look at too for exchange? Josh to discuss with Lars Boehme and Xavier Hoenner
* **ACTION** In sync with this exploration - manufacturers (and maybe other users) could think about what the “rawest” form of each data stream is and how to record the programmers choices and any analytical thresholds (e.g. how a ‘dive’ is defined: depth/time/wet or dry, etc.)
* Ed Bryant raised the question about how raw is “raw”? For example the rawest form of satellite data doesn’t even have locations. What about accelerometry data? Something for the working group to think about further.
* Josh London– is it time for us as a community to look beyond a .csv or “flat file format”? For example JSON or xml? Lars Boehme suggested that this is already happening with– multi-dimensional arrays in netCDF for example.
* **ACTION** Task team to discuss fie formats (Hoenner, Boehme, McIntyre, Bajona, Roquet- see below)

Data archiving

* Melinda Holland – manufacturers are not guaranteeing an indefinite archive, at least Wildlife Computers is not guaranteeing an archive.
* There was a suggestion to start with TDR data first then location/position data then perhaps summarized data for developing the standard? (NB location data are a time series too). But much of the data under discussion is not TDR and the data exchange may well start with other formats
* Melinda also noted that for archive data that WC currently has they would need some resources ($$) to transition/translate that into whatever format we want.
* Lars Boehme– pointed to Argo as a good example we could learn a lot from.
* Sam Simmons: Practical implementation for translating to a new “standard” – how far back in time do we want to go? Or do we just start from now with the firmware metadata translation, etc? Not sure this was resolved but something for everyone to be thinking about.
* **ACTION** Users and manufacturers following this meeting will work on/provide lists for what they each feel are important to preserve in the metadata/data records.

**Animal Borne Task Team**

* Lars Boehme: formation within **EOOS of an “animal-borne task team”** (so far that consists of Lars and Fabien Roquet) but currently are tagging animals on the European shelf and there are connections with MEOP internationally. Currently the tag manufacturers are working on Matlab/R code to get tag data into netCDF and they are focused on oceanographic data and making that available initially, but want to deal with behavioral data eventually.
* **ACTION** Hoenner, Boehme, McIntyre, Bajona, Roquet to connect on this point and discuss the AAT standard (I think OTN using R only at this time not Matlab).
* **ACTION** Peter Walsh– suggested for the next step we focus on the AST draft document and provide comments and feedback on what we like/don’t like and (.and think about what the bigger picture is: where do we want to go, what are the mechanisms for sharing/requirements for compliance, etc.