OTN Canada Scientific Advisory Committee (SAC) Meeting
Tuesday-Wednesday, 3-4 December 2013

Location: Somervell Meeting Room, Third Floor, Delta Vancouver Suites, 550 West Hastings Street, Vancouver, Canada

Attendees:

- In person: Chris Barnes, Nikki Beauchamp, Steve Cooke (*Chair*), Sara Iverson, Katja Fennel, Ian Fleming, Aaron Fisk, Michelle Heupel, Scott Hinch, Alison Janidlo, Kyle McKenzie, Kes Morton, Dale Webber, Fred Whoriskey

- By teleconference: Bob Branton (invited guest; OTN Data Centre), Tracy Rounds, Alain Vezina

• Regrets: Svein Vagle

Data Management and Products

(text to accompany Intelligent Open Access presentation)

Bob Branton presented on data management contributions to OTN Canada, among which are the OTN Data Explorer or Atlas project, as well as intelligent open access. OTN data policies are aligned with data standards and best practices of international funding agencies. To remain current, OTN periodically reviews policies against emerging expectations. The initial policies under which OTN was set up were the OECD recommendations. In 2011 there was the Canadian Research Data Summit, where it was announced that any data from Canadian publicly funded research should be available as soon as possible. After looking at our policy, the feedback we gave through CODATA based on biologists' expected exceptions, was we wouldn't report on endangered species or for HQP in training. This is what we refer to as intelligent open access.

Last year, we planned a data management workshop, which was held in June with the symposium, where policy and tools such as Google flyover and public download were reviewed. Additionally, last June a brand new feature, private detection extracts, was rolled out where individual researchers could share detailed detection data for their animals, regardless of where it was detected, and they had control over who saw that. We went over all the forms and templates that OTN needs to have filled in, and then into the common holdups and how to avoid them. We held a user/manager panel, for new product suggestions, with Ian Jonsen, Eddie Halfyard, and Mathieu Dever on the panel, and Diego Ibarra was the moderator.

When we first built the OTN data system, we only had deployments, and not much in the way of tracking data. Now we actually have lots of detections, 66 species, 35k animals on a worldwide basis. Now it is time to develop products a tracker would like to see, a map of the release tags, green where they have been detected, and the dash with detection counts, in log-10 scale. The data would be embargoed, or held private, for two years and then released. We would negotiate with people to make their data public on a date, and then commit to something that fits to them.

Public data will be published to OBIS, where OTN will ultimately contribute about 100 collections. Right now there are six collections, including the Inner Bay of Fundy salmon, NS power project, Sable Island grey seals and a couple of others. We're tracking the referrals through OBIS. We're publishing metadata for the arctic to the polar data catalog, which is reviewed with Aaron, and that should be on the polar data system in time for ArcticNet next week.

We are mapping OTN detections with receiver metadata online. When you click on the release project, it pops up with a link to the OTN release metadata, WORMS information, as well as, if it has been released to OBIS, the link to OBIS. This is a clearinghouse, a visual way of getting our data, which is quite interesting. Another big breakthrough was with Diego Ibarra in the Dalhousie Oceanography Department. We gave him links to the system in St. Andrews, where grey seal people do their data preparation. This too is linked to the discovery metadata and tracked through Google analytics, so we know people are using our system.

One of the things that came up in the last year, in dealing with all this CODATA summit, people are wondering what to do with non-tracking data that don't originate with tags. OTN and CHONe are working on a system affiliated with published papers that are put into a system called DRYAD, can put several data packages into a data set and get digital object identifiers to the paper, and embargo the data, to deal with idiosyncratic data unique to particular projects. As well, DRYAD expects the data to published, while the big share allows the same kind of thing with unpublished data. We don't have to build a system like the data warehouse to manage these data. Others are presenting others kinds of slides, the preferred way to deal with smaller datasets.

Other developments are underway, in particular to tracking, sentinel tags, and to anonymize detections so that a line operator could actually do some analyses. Now we have 66 million records in our system. Tsafrir Gazit and Richard Apostle published a paper in the *Journal of International Wildlife Law & Policy*, dealing with the issues of intellectual property rights, which is a lead in to a CFI application with the Dal Librarian for the 2015 innovation fund, regarding research data management towards an integrated institutional framework. It is at the LOI stage, and the librarian has accepted it as a submission.

OTN had a meeting with Hydra, who have the primary data management project in the Puget Sound area, and we are going to routinely exchange data with Hydra, metadata and such, and starting with the green sturgeon project run by Phil Dion and company. The idea is that any tracker, if they follow a policy, can identify themselves and look at mystery tags and identify their tags, then OTN will set them up right away, give them what they have now and anything in the future. Susan Dufault is the main contact for that.

Data Embargo

As a data manager, we want the data into the system, secure and make sure it exists, and then leave it up to the researchers to decide which data become public. They need to commit to making data public. If OTN Canada as a group wants to collectively define the embargo period, we have set it at two years, and anything beyond that is an exception. However, if you are too strict, trackers will hold the data. For a detected tag, there are on average 2000 detections for each tag. We have to be gentle with the trackers to give the details of their tags to the system. If they withhold, we have nothing. It's a balancing act, takes diplomacy, have to be in the system and public.

Those who deploy equipment and release tags also receive funding from other sources. Only 15% of the detections are from CFI or NSERC funded projects. We don't own the tags, the people giving the data to us, are giving us with the idea of our policy in the background. They are committed to making the data public, but with a tag in particular, if we were to release half the record of the tag and make that public, and the researcher is trying to get the full picture, they will interpret half or three quarters of the story. We want the whole story on the tag.

We're trying to generate useful and interesting products that are public. If an animal is detected by a glider, we can make that information public and do it in such a way that the trackers want us to do that for them. We're encouraging that to be public. Same with the mapping and things like that. We're not doing that for the individual but for the collective. To have them appear, they need to be made public sooner than later.