Case 1: Professor Periwinkle

Professor Periwinkle leads a large team of researchers, students, and staff. Her team tracks marine wildlife using a variety of sensors and monitoring equipment.

Data is collected digitally by remotely-operated marine vehicles (ROMV), tags that are surgically implanted in captured and released animals, static sensor buoys that measure ocean conditions, and communication lines that passively listen for signals from animal tags. Each collector produces data in a format specific to the sensor, but Dr. Periwinkle's team has developed software that converts the data to the NetCDF format for use and storage; the sensors are currently producing about 300 MB per day in raw sensor data which becomes 500 MB in uncompressed NetCDF formatted data (both versions are kept). Dr. Periwinkle's website also allows citizen scientists to report sightings of wildlife in the Minas Basin, though this data is not currently used. Other forms of data include field notes that document the animals captured and tagged; field notes from mark-recapture population estimation experiments; and field notes from observational studies conducted during classes Dr. Periwinkle teaches. Finally, two members of Dr. Periwinkle's team run complex simulation models that attempt to predict animal populations and movements based on oceanographic data acquired from the buoys and from collaborator; these models produce gigabytes of data in zipped .csv files.

The data collected is used to write papers and to create exercises for graduate classes.

Data is made available in visual form to the general public using OceanViewer.org, a platform for sharing oceans-related data.

There is currently no structured sharing of data outside of the university but the long-term goal is to make this data available for use; after Dr. Periwinkle writes a paper on the impact of ocean acidification on seal populations with her data, someone else might want to look at other chemical processes or other animal populations with the same data. Currently they can email Dr. Periwinkle to request the data files, and Dr. Periwinkle can upload them to Dropbox. (The university's file sharing provider, Microsoft

OneDrive, requires external users to login with a Microsoft-provided account of some kind, and not all people have these accounts).

There are other research teams in Canada and around the world that collect data that might be interesting to include in Dr. Periwinkle's team's research; they currently request access to this data through personal connections and relationships. Organizations in Canada collect similar data, like OTN at Dalhousie, ONC in BC, SLGO in Quebec, SmartAtlantic in Newfoundland, ArcticNet, the department of fisheries and oceans, Environment Canada. There are also individual research groups collecting data in specific species, specific sites, and more. There is not currently a data sharing agreement in place, merely a collection of ad hoc sharing opportunities. In contrast, NOAA in the US has the IOOS system which spans the country, and the European Commission has EMODnet.

Everyone who starts with Dr. Periwinkle's team is told how the data is formatted and structured, and data is primarily shared within the team using external hard drives and USB keys. There is a help document from 2003 describing how to work with the data sets, and how to contribute to the datasets; this help document is on a shared drive somewhere. As students graduate they share their data with current students in the lab, or take it with them on their personal devices.

Data has been collected in some capacity in Dr. Periwinkle's lab since 1998; over the last three years, however, her funding has increased and she is now collecting substantially more data than ever before. Dr. Periwinkle is proud of a shelf in her office which contains all of her past data in various formats, reflecting the evolution of technology: floppy disks, ZIP Disks, CDs, DVDs, BluRay, and external hard drives.