

Data Management Plan

The projects conducted under this data management plan include all projects managed by Coonamessett Farm Foundation (CFF). The data generated includes oceanographic, weather, catch, behavioral, physiological, demographic and operational efficiency measurements. Data collection following this data management plan began in 2010 and covers Georges Bank and the Mid-Atlantic. Data types captured include digital numeric data, photographs, video, database tables, spreadsheets, paper records, physical samples, etc. The data is captured/ created using satellite, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle (AUV), remotely operated vehicle (ROV), animal tagging, surveys, etc. The data management described here will be stored and displayed on the CFF website, the CFF database and the CFF server. None of the data collected will contain personally identifiable information or any information whose distribution may be restricted by law or national security.

Point of contact for this data management plan is:

Name: Mary Newton Lima
Title: Administrative Assistant
Location: Coonamessett Farm Foundation Inc.
Email: mary@cfarm.org
Phone: (508) 356-3601
Address: 277 Hatchville Road
East Falmouth, MA 02536

Mary Newton Lima is the point of contact for the data collection and will ensure that questions are redirected to the appropriate Principal Investigator (PI) for each project. The PI of the project is responsible for verifying the quality of the data and answering specific questions about the data collection. Jess Holderbaum is responsible for the data storage and data disaster recovery activities. Ms. Newton Lima is also responsible for ensuring adherence to this data management plan, including ensuring that appropriate resources are available to implement the data management plan.

Data is collected and entered into the CFF database by the PI using the follow protocols:

- Catch data is identified under a trip identification number (ID) that consists of a three letter vessel id, a three digit trip number, and two digits indicating the funding year (ex. a trip on F/V Celtic that was the 89th trip taken by CFF and was receiving funding from 2012 the trip ID would be CEL-089-12).
- Picture data is saved directly to the server under a picture ID comprised of a three digit trip number and a three digit tow number. If the project analyzes catch from each side separately, a 'port' or 'stbd' label is added to the end of the picture ID to indicate port or starboard (ex. for a photo taken of the catch on the starboard side of the vessel of the catch collected after tow number 22 on the 108th CFF trip, the picture ID would be 108.022_stbd).
- Video data is saved directly to the server under a video ID according to the equipment being used. ROV and autonomous underwater video AUV video IDs consist of the date (year month day) and the time stamp. If a sea turtle is captured in the video, a running turtle number will be added to the end of the ID (ex. for a ROV video taken of Turtle 27

on June, 17th 2008 at 11:08am, the video ID would be 2008 06 17 1108 Turtle 27). All other video IDs consist of a three digit trip number and a three digit tow number (ex. for the 76th trip taken by CFF during tow number 68 the video ID would be 076.068).

Following data input into the database the database administrator performs an initial quality control check on the data using basic queries developed to identify common errors. A summary of errors identified is then sent to the PI. The PI then corrects any errors, verifies the data and compiles a trip summary from the data. The data and trip summary are then saved to a locked version of the database that is held on the CFF server and backed-up on external hard drives off site.

The CFF file server will be used as a data repository and to document the data collection. In addition to raw data, other documentation such as final reports and presentations will fully describe the data and ensure its long-term usefulness. All geographic data will meet ISO 19115 standards (Geographic Information – Metadata).

Data is expected to be made available to the public within a year of project completion. There may be a PI hold or other delay between data collection and publication that will take no longer than three months. Users of the data will be subject to access conditions or restrictions through a data request process that includes submission of nondisclosure statements and special authorization from that PI or a PI representative. An open-standard, interoperable, non-proprietary web service such as a secure ftp will be managed on site to enable data sharing.

Data will be stored to the CFF database and the CFF server before being sent to a long-term archive facility. The data will be protected from accidental or malicious modification or deletion by storing all data within external hard drives stored. All data will receive daily onsite back-ups from the database to the server and weekly off-site back-ups onto external hard drives. All data on the CFF database and server is protected from unauthorized access using encryption and industry standard security measures. In the event of unauthorized access the security breach will be identified and closed immediately. The source of the security breach will be investigated and a change in security measures may be considered. Individuals that use the data set will be notified that data may be compromised due to unauthorized access.

The data will be archived and preserved in the NODC and NCDC NOAA Data Center. Our long-term strategy for maintaining, curating, and archiving data outside of the NOAA Data Center are the CFF database and server maintained on site. Costs of long-term data archiving will be provided and maintained through awarded grant funding.

All transformations or procedures necessary to prepare data for preservation or sharing will be completed by the PI including but not limited to: quality control, format conversion, and anonymization of personally-identifiable information. Proper indexes will be created upon completion of the project for related information such as but not limited to: references, reports, research papers, algorithms, audio or video codecs, and special character sets or fonts to enable future use and understanding of the data upon submittal to the archive.