NOAA CORAL REEF CONSERVATION PROGRAM

Coral Data Management Planning

June 20, 2011

Jeffrey Campbell, Ph.D.

jeff.campbell@noaa.gov

NODC's 50th Anniversary 1960 - 2010

ERT

Imagine

Imagine how much better and more productive your research, monitoring, and resource management would be if you had extensive baseline data from 50 years ago.

- How could you learn the data exists?
- How could you obtain the data?
- How would you know if the data are suitable for your new, unanticipated use?

Agenda

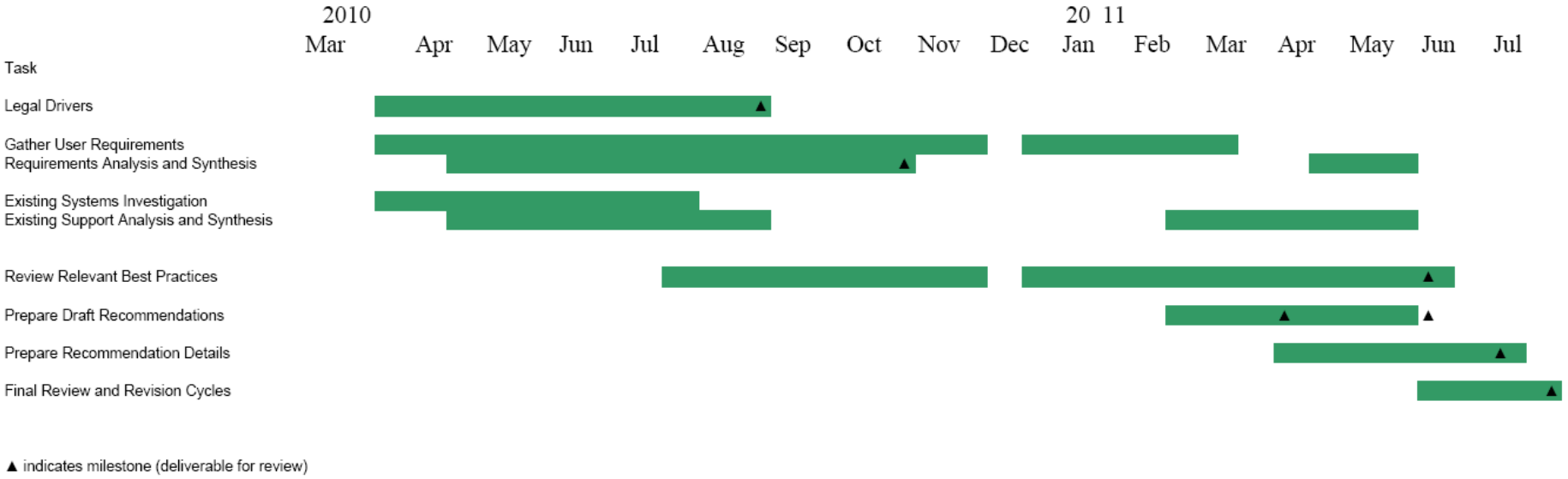
- Project Introduction
- Data Management Plan Overview
- Lessons Learned

Coral data direct collection via USB cable

Introduction - Project Resources

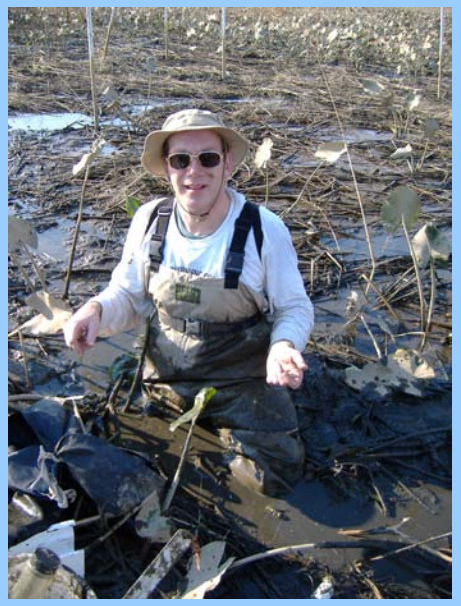
- Consultant – full time
- Coordinator – 2 hours/week
- Management support for:
  - subject matter expert interviews (within CRCP, NNDC, PIs, etc)
  - Timely review and comments on drafts
- Mid-management guidance
- Document Review Committee (17)

Introduction - Actual Time Frame



Consultant Introduction

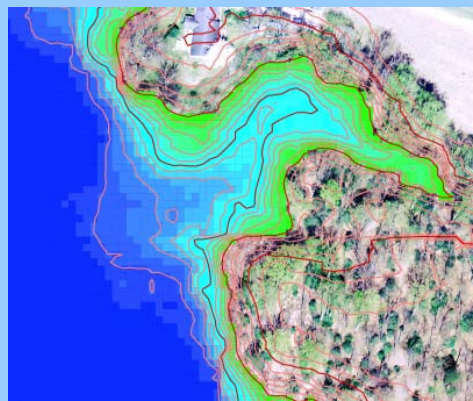
- PhD Information Science – University of Pittsburgh (Pitt)
- MSIA - Carnegie Mellon University (CMU)
- BA Geology – University of Rochester (UofR)
- Consulting
  - Manager, Accenture (formerly MICD AA&Co)
  - Campbell Systems Consulting (CSC)
  - Environmental Informatics (EI) www.EnvironmentalInformatics.com
- Teaching
  - Assistant Professor, Information Systems, UMBC
    - Systems Analysis and Design
    - Database Management
    - Intro to Human Computer Interaction (HCI)
    - Computer Supported Cooperative Work (CSCW)
  - Carnegie Mellon, Carlow University, Univ of Pittsburgh
  - University of Pittsburgh Computer Leaning Center



Obligatory field photo

Consultant Introduction

- Research Scientist, Center for Urban Environmental Education (CUERE), UMBC
- Scientific Advisory Board, Jug Bay Wetlands Sanctuary (JBWS), Chesapeake Bay - National Estuary Research Reserve Maryland (CBNERR-MD)
- **Information Systems Research**
  - HCI (making systems easier to use and more effective in communicating information)
  - CSCW (remote support for collaborative work) – diagrams, text, GIS
- **Environmental**
  - Real-time collection and visualization of geospatial hydrologic data
  - Ecological Data Warehouse for NERRS site integrating local and USGS, NERRS, NCDC, MD Dept of Natural Resources, US Naval Observatory data
  - Data Management Committee Chair – MD Water Monitoring Council (MWMC)
  - GIS analysis for NERRS site
  - Internet watershed mapping for watershed management plan



Introduction - Consultant's Approach

- Independent, unbiased
- Client oriented
- Confidentiality
- User-centric focus
  - "User" broadly defined
  - What do the users truly need?
  - What is the best way to meet those needs?
- Larger context
  - State of the art, best practices, standards, etc.

Users

National Research Council

"… all environmental data management activities including … the development of the enterprise-wide data management plan, should incorporate substantial and ongoing user input." (page 3).

CRCP DMP Framework

The user types include other NOAA programs/offices; federal, state, territorial and local agencies; domestic and foreign universities; fishery management councils; domestic and international NGOs; local communities; the general public; and foreign governments. (page 18)

Coral Reef Conservation Act of 2000

- preserve, sustain, and restore coral reef ecosystems
- promote wise management and sustainable use
- develop sound scientific information, and
- support local communities

2009 National Goals and Objectives

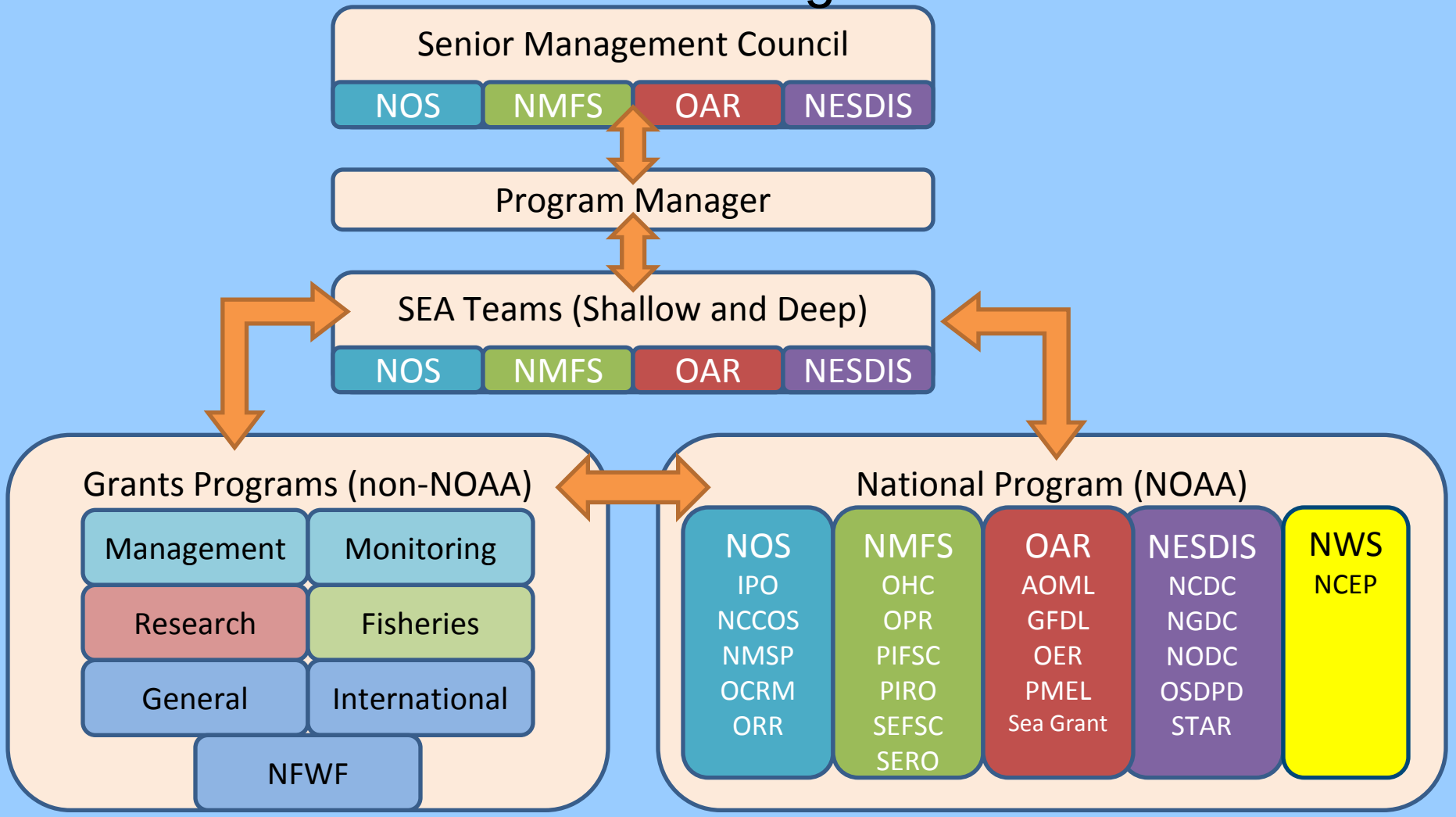
Climate Change

Fishing

Land-Based Pollution

NOAA CORAL REEF CONSERVATION PROGRAM

CRCP Matrix Organization



CoRIS

Coral Reef Information System
coris.noaa.gov

Red Coris Wrasse
Coris gaimard

- Track, record and report delivery of data and information deliverables for CRCP funded projects
- Develop, maintain and provide access to data and information deliverables through a metadata-driven web site
- Develop and maintain CoRIS web site capability
- Provide metadata training and guidance to CRCP project teams

Data Management Plan Overview

Shark guarding data warehouse

Scope of Data

What data is to be included in the DMP?

- All data with any CRCP support:
- Real-time, periodic, research, experiment
- Video/photos as primary data
- Biological data
- Socio-economic data
- Mostly geospatial

Data Management Functions

CRCP Environmental Data Management Functions and Life Cycle								
Planning	**Observation Acquisition & Transmission**		**Scientific Data Management**			**Archive & Access**		
Observing Capabilities & Requirements	Maintain & Monitor	Collect & Rescue	Calibrate & Validate	Appropriate Formats	Complete Metadata	Long-term Preserv-ation	Data Discovery	Access / Disseminate
Contingency Planning								

Project Deliverables

- Data Management Report
  - Detailed analysis and recommendations
- Data Management Summary
  - Extended abstract of Data Management Report
- Coral Data Policy
  - Recommendations primarily for CRCP HQ
- Data Management Plan Template
  - Project level plan to be prepared during proposal and award phases

Data Management Report

Planning	**Observation Acquisition & Transmission**		**Scientific Data Management**			**Archive & Access**		
Observing Capabilities & Requirements	Maintain & Monitor	Collect & Rescue	Calibrate & Validate	Appropriate Formats	Complete Metadata	Long-term Preserv-ation	Data Discovery	Access / Disseminate

- Legal Drivers
- Functional Requirements
- Current Systems
- Best Practices
- Recommendations

1. DMP Introduction
2. Data Management Functions
    1. Planning
        1. Drivers
        2. Functional Requirements
        3. Current Systems
        4. Best Practices
        5. Recommendations
    2. Data Acquisition and Transmission
        1. Maintain & Monitor
            1. Legal Drivers
            2. Functional Requirements
            3. Current Systems
            4. Best Practices
            5. Recommendations
        2. Collect & Rescue
            1. Functional Drivers
            2. User Requirements
            3. Current Systems
            4. Best Practices
            5. Recommendations
3. Scientific Data Management
    1. Calibrate & Validate
        1. Legal Drivers
        2. Functional Requirements
        3. Current Systems
        4. Best Practices
        5. Recommendations
        6. …..

CRCP Data Policy

- Summary and Recommendations for each stage of the life cycle
- Policy Recommendations for
  - NOAA-level issues
  - Data discovery enhancement strategies
  - Data provider benefit and cost discussion
  - Data management organization change
  - Data management evaluation

Data Management Plan

- Project-level
- Questions to guide responses

- DMP required details increases from pre-proposal to award

Data Management Tiered Approach

- Federal requirements
- NOAA policies
- Line Office policy
- CRCP data policy
- Project group policy
- Project-level data management plan

Lessons Learned

- Report structure works
- Review cycle “cliff hangers”
- Integration of science and IT is crucial
- Additional support for data providers
  - Reduce effort (cost)
  - Increase benefits
- Enhanced Discovery and Access for data consumers
- Issues beyond scope of the program

Lessons Learned – Broader Issues

- Archive for still and moving images as source data
- Data apparently outside NNDC scope
- Socioeconomic data
  - Qualitative research, ethnographic work, and focus group transcripts produce significantly different data than biological/physical values.

Jeffrey Campbell, Ph.D.

jeff.campbell@noaa.gov

301.713.3272 x119

Full Data Life Cycle

- Planning
- Observation Acquisition & Transmission
  - Maintain & Monitor
  - Collect & Rescue
- Scientific Data Management
  - Calibrate & Validate
  - Appropriate Formats
  - Complete Metadata
- Archive & Access
  - Long-term Preservation
  - Data Discovery
  - Access/Disseminate
- Contingency Planning

Project Phases

- Legal Drivers
- Functional Requirements
- Current Systems
- Best Practices
- Recommendations

Significant Aspects

- CRCP spans Line Offices
- Geospatial biological and physical data
- Leading edge for NOAA project-level data planning


jeff.campbell@noaa.gov

Best Practice Topics

- Metadata
  - CSDGM vs ISO
  - Quality metrics
  - Designated community standards
- Other relevant standards
  - OAIS
  - OGC

Best Practice Topics

- “Difficult Data Types”
  - Video, still photos as original data
  - Human dimensions (socioeconomic data)
- Data Discovery
  - Precision and recall
  - Semantic search, ontologies, controlled vocabulary
  - Geoportal vs Google Earth/Maps
  - Federal portals (data.gov), web services

Requirements

- Systems Analysis and Design analogy
- Must identify the infrequent and unusual requirements, even ad hoc and emergency needs
- Managerial requirements vs data requirements vs data management requirements
- Not specific data elements but
  - types (video, image, GIS layer, etc)
  - related data / metadata
  - retrieval characteristics (search criteria)
  - Change over time…

Requirements Gathering Themes

- Broad Perspective
- Data Focused
- Resource Management

Requirements Gathering Themes

Broad Perspective

- Current data archiving techniques
- Variation of data types year to year (within a project)
- New tools/products to help you manage data
- New tools/products to help you accomplish mission
- Other relevant data management planning

Requirements Gathering Themes

Data Focused

- Types of data
- Sources of data
- Types of data most difficult to manage/archive
- Monitoring of automated sensors
- Calibration/validation processes/issues; external data usage?
- Factors changing observations
- Non-computerized data
- Extent of modeling, nature of modeling data
- GIS capabilities

Requirements Gathering Themes

Resource Management

- Current sources of information for decision making
- Relevant non-NOAA data (sources, format)
- Data provided by you to citizens, NGOs, government, etc
- Extent to which volunteers collect data