

Video Data Management System Archives and Provides Online Access to NOAA Deep-Sea Corals Digital Video and Image Data

A. Fiolek

NOAA Central Library
SSMC3, 2ND Floor, 1315 East West Highway
Silver Spring, MD 20910 USA

D. W. Collins

NOAA National Oceanographic Data Center
SSMC3, 4th Floor, 1315 East West Highway
Silver Spring, MD 20910 USA

Abstract- Since late 2002, the National Oceanic and Atmospheric Administration's (NOAA) Office of Ocean Exploration and Research (OER) has been collaborating with the National Oceanographic Data Center (NODC) and its three divisions, the NOAA Central Library (NCL), the Marine Data Stewardship Division (MDS), and the National Coastal Data Development Center (NCDDC) to address the requirements for archiving, preserving, managing, and providing online access to digital videos and still images from OER oceanographic expeditions.

The Video Data Management System (VDMS), which was developed to facilitate managing online information and access to video and images obtained during NOAA-sponsored oceanographic expeditions, now enables access to hundreds of digital video clips, highlight movies, still images and related documents and products from OER expeditions. Using discovery tools available via NOAALINC, the NCL online catalog (<http://www.lib.noaa.gov/>), NOAA scientists and other researchers can discover and download online video and still images to deep sea coral ecosystems areas. Upon special arrangement, the NCL can provide equipment and an appropriate environment for users to view, copy and/or download requested off-line scientific video data or view original expedition tapes from the NOAA Library Archives. Additional online information includes related cruise reports, educational lesson plans, original video and image annotation documents, digital maps and Web sites, and links to other oceanographic observation data.

During the development of the VDMS project plan, which is a part of a larger comprehensive OER Data Management Project, the VDMS team defined and established several 'best practices' to support OER video data management requirements. Metadata guidelines for digital video (DV12) and digital still images (DI12) help scientists and data managers in the field to create complete descriptive metadata about their image data. Scientists, librarians and archivists then use this information to create MARC21, FGDC, or Dublin Core metadata records. The VDMS team also developed a work-flow for managing digital video by defining the process for moving video data from ship to library to archive, including steps for creating archival backup copies and web-accessible video clips and highlights.

The VDMS presently manages off-line access to more than 1500 MiniDV and 500 DVCAM tapes, over 1500 DVDs, and online access to more than 300 digital video clips and highlights collected during NOAA ocean exploration cruises. Over 80% of all digital video and image data are from OER expeditions to various deep-sea coral areas. A growing collection of the digital data, including *in situ* physical and chemical ocean observations are archived in NCL and NODC. *In situ* data are accessible through the search and retrieval functions of the NODC Ocean Archive System (OAS) at

<http://www.nodc.noaa.gov/Archive/Search/>.

I. INTRODUCTION

Many types of data and information products are collected and created during a typical oceanographic cruise, including planning documents, cruise summary reports, laboratory specimen lists, video and still images, and navigation and other observational data. As required by NOAA Administrative Orders NAO 15-217 and NAO 205-17, the NOAA Central Library (NCL) and National Oceanographic Data Center (NODC) are receiving an increasing number of video data collections from diverse NOAA components, including the Office of Ocean Exploration and Research (OER), the National Marine Sanctuaries Program (NMSP), and the Coral Reef Conservation Program (CRCP).

Beginning in late 2002, the NCL and NODC began collaborating with OER data managers to develop and implement an end-to-end data management plan for data and information collected during OER-sponsored cruises. The Integrated Product Team (IPT) was formed to develop a comprehensive plan, with several working groups focusing on each component of the overall plan. One working group was tasked with developing a Video Data Management System (VDMS) for acquiring, cataloging, archiving, maintaining and providing access to digital video data. The IPT also recognized the added benefits of developing VDMS requirements, documentation and processes that could serve as a model for all NOAA scientific video data.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE SEP 2008		2. REPORT TYPE		3. DATES COVERED 00-00-2008 to 00-00-2008	
4. TITLE AND SUBTITLE Video Data Management System Archives and Provides Online Access to NOAA Deep-Sea Corals Digital Video and Image Data				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Oceanic and Atmospheric Administration (NOAA), Central Library, SSMC3, 2ND Floor, 1315 East West Highway, Silver Spring, MD, 20910				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM002176. Presented at the MTS/IEEE Oceans 2008 Conference and Exhibition held in Quebec City, Canada on 15-18 September 2008.					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

This paper describes the processes to assure that digital video data from the underwater explorations are archived for the long term and managed consistently and effectively with minimal staff resource requirements.

The VDMS objectives are:

1. To provide timely online access to OER video and still image data.
2. To educate the public about NOAA oceanographic expeditions and underwater explorations through related lesson plans.
3. To archive and preserve unique video and related data for the long term.
4. To foster collaboration between NOAA librarians, data managers, and scientists from different line and program offices.
5. To use or extend existing library tools, guidelines, and metadata standards to support new media formats, such as digital video, digital image, and digital text documents.
6. To enhance data access and metadata sharing between the NCL's NOAALINC, NODC's Ocean Archive System (OAS) and NCDCC's MERMAid catalogs, CoRIS and OER Digital Atlas databases.

II. VIDEO AND IMAGE DATA MANAGEMENT

The primary media currently used for capturing video images are MiniDV and DVCAM digital tapes or DVD-R discs. Each of these media types uses different encoding (file format structure). Video processing software is required to convert the native uncompressed video formats to current industry and archival standard formats (i.e., uncompressed DV or AVI) to facilitate online access and long-term management. At present, the VDMS process uses DV and/or AVI formats as the archival video encoding and MP4 and MOV as the online access encoding. Video files intended for online access are encoded with MPEG-4 and/or H.264 compression codecs which are supported by widely used video players, such as QuickTime™, Windows Media Player™, and RealMedia™.

At present, the NCL manages a growing collection of multiple video media. This collection includes more than 1500 MiniDV tapes, 500 DVCAM tapes, approximately 1000 VHS tapes and more than 1500 DVDs. Over 80% of the collection are videos from deep corals ecosystems in the Gulf of Alaska, Gulf of Mexico, along the North American Atlantic and Pacific coasts, and other locations. These original media contain the entire sequence of video footage captured with standard resolution video equipment during dozens of cruises. In some cases, the video was captured using high-definition (HD) video equipment. These videos provide a relatively complete record of events, submersible dive or related activity. Original video media are currently stored in a climate controlled room and are being migrated to new mass-storage media for ongoing long-term archival preservation and improved access.

In addition to original media, the NCL archives, and provides online access to clips and highlights created from the

original, full length raw video. Clips typically contain very short (10-60 seconds) excerpts of interesting or unusual features. Highlights are usually a series of short video segments (2-15 minutes) selected by the principal Investigators (PI) and/or data manager as a representative sample of images collected during the cruise.

Figure 1 shows a Quick Time™ frame from the highlights video from the *Life on the Edge 2005* expedition to a deep corals bank in the South Atlantic Bight.



Figure 1. Quick Time™ video frame from highlights of the *Life on the Edge 2005* expedition investigating deep corals of the South Atlantic Bight.

The video highlight on MiniDV tape was “captured” using Apple™ Final Cut Pro™ (FCP) video editing software. The captured file was then exported to DV and AVI archival formats using FCP Compressor™ and exported to MOV and MP4 formats for the online accessibility. By testing different codecs and file format combinations, the VDMS team developed a Standard Operating Procedure (SOP) for in-house digital video conversion. The VDMS SOP uses a ‘best practices’ approach, based on the limited guidelines for managing digital video processing and conversion provided by national authorities, such as the Library of Congress and the National Archives and Records Administration.

III. ONLINE ACCESS TO DEEP-SEA CORAL VIDEO AND STILLS

The NOAA Central Library provides online access to clips and highlights via links in standard MARC21 records in the library online catalog – NOAALINC (<http://www.lib.noaa.gov/>). A search in NOAALINC for “deep-sea corals AND digital video” returns a list of all catalog records that include links to the digital videos on deep-sea corals, as well as links to other related media and documents. Figure 2 shows an example of the NCL metadata record in MARC21 standard for the *Expedition to the Deep Slope 2007* investigating deep-sea coral areas of the Gulf of Mexico.

The screenshot shows the NOAA Library and Information Network Catalog interface. At the top, there's a search bar and navigation links. Below, a search result for "Expedition to the Deep Slope 2007" is displayed. The record includes a title, electronic access information, related links, and a detailed description of the collection. The description mentions that the collection includes 630 DVD-ROMs, ca. 1260 hours of digital video data, and 19 digital video highlights on MinDV tape. It also notes that the key to this kind of expedition is a remotely operated vehicle (ROV) system named Jason II.

Figure 2. Example of NOAA LINC metadata record in MARC21 standard format. Collection-level record describes Deep Slope 2007 expedition.

Currently, there are 28 digital video collections in NOAA LINC using the criteria listed above. These links provide off-line access to over 1500 original video tapes. The links also provide online access to more than 300 clips and highlights, as well as over 6000 still images. Some additional examples of library metadata to NOAA Ocean Exploration and Research cruises to deep coral ecosystems include:

- Gulf of Alaska 2002 - Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocn232154662>
- Gulf of Mexico Deep Sea Habitats 2003 – Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocn232118892>
- Gulf of Alaska 2004 - Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocn232118892>
- Estuary to the Abyss 2004 - Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocm63173772>
- Life on the Edge 2005 - Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocm70078821>
- Investigating Charleston Bump 2003 - Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocm57306431>
- Operation Deep Scope 2005 – Collection
<http://www.lib.noaa.gov/uhtbin/cgiirsi/x/x/0/5/?searchdata1=ocm63675897>

More information about NOAA deep-sea coral video, image, and scientific observations from oceanographic explorations is also available using NOAA's Coral Reef Information System (CoRIS, <http://www.coris.noaa.gov/>), the OER Digital Atlas developed by the NCDCC at http://www.ncddc.noaa.gov/website/google_maps/mapsOE.htm, and the NODC Ocean Archive System (OAS, <http://www.nodc.noaa.gov/Archive/Search/>).

An accurate inventory of original still images, usually copied from separate CD-ROMs or DVDs, is more difficult to obtain and organize. For example, a single collection of still images captured during one OER cruise includes more than 90,000 images. However, many selected still images of deep-sea corals and coral reef animals (see Figure 3) are available through the online NOAA Photo Library, a digital image collection of over 38,000 photos. The collection is arranged

into subject matter albums, including *Coral Kingdom* (<http://www.photolib.noaa.gov/reef/>), *Caribbean Seas* (<http://www.photolib.noaa.gov/reef/caribbean.html>), *Florida Reef* (<http://www.photolib.noaa.gov/reef/caribbean.html>), *Pacific Reefs* (<http://www.photolib.noaa.gov/reef/pacific.html>), and *Red Sea Reefs* (<http://www.photolib.noaa.gov/reef/redsea.html>).



Figure 3. Free floating giant Caribbean anemone - NOAA Photo Library. View HR image at <http://www.photolib.noaa.gov/bigs/reef1081.jpg>

IV. DOCUMENTATION AND METADATA MANAGEMENT AND EXCHANGE

Many types of metadata are required to ensure that digital video and still images can be interpreted and used by software that may be available on future computers and operating systems. Technical metadata about the encoding format, codecs used, playback rates, etc., are critical for software to be able to correctly interpret and display the contents of the digital files. Descriptive metadata is necessary to provide context for the content being displayed and answers questions about when, where, and why the images were obtained and who obtained them. It is often easiest to acquire these types of descriptive and technical metadata from the PI or data manager soon after each cruise.

Video shot during submersible operations is often the primary data collection activity for the dive and is intrinsically a form of geospatial data. As a result, the use of video as a source of quantifiable geospatial data (e.g., percentage of seafloor area covered by sponges and echinoderms, identification of common and unique species at a specific location) makes the content management and metadata requirements different than video that is primarily a record of a historic event.

The VDMS working group developed a set of tools and documents to assure standardized access to data and information collected during NOAA OER-sponsored cruises. These include:

- A VDMS requirements document defines technical standards for the system, describes archival storage formats and conditions, and specifies online retrieval requirements.

- Metadata standards requirements include using the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata and MARC21, the international library standard. Other systems, such as CoRIS, may access VDMS information in NOAAALINC using the Z39.50 search protocol.
- A crosswalk and converter in the MERMAid system (NCDDC online catalog) enables sharing common metadata in both FGDC and MARC21 metadata records, as well as converting FGDC metadata to MARCXML format.

Collection-level (parent) MARC21 metadata records from the NCL can be linked to dive-level (child) FGDC metadata records created by the NCDDC using the crosswalk and conversion techniques developed for the VDMS. These metadata records provide the descriptive information for both NOAAALINC and MERMAid metadata discovery tools (Figure 4).

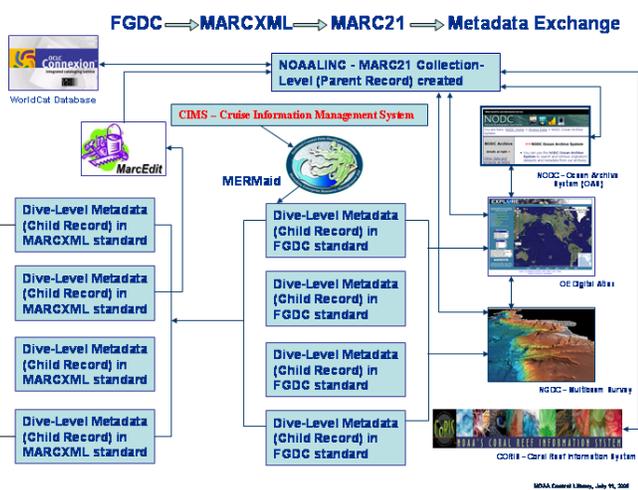


Figure 4. FGDC-MARCXML-MARC21 Metadata conversion and exchange workflow

http://docs.lib.noaa.gov/OEDV/VDMS_DOCS/MARCXML2008.pdf

- Each collection, tape, clip or highlights video, or still image may be documented using guidelines developed by the VDMS working group. These guidelines are referred to as DV12 (Digital Video 12 descriptive elements) for video and DI12 (Digital Image 12 descriptive elements) for still images. DV12 and DI12 templates were developed to help field personnel document the contents of videos they created. Information from the field personnel (in DV12 or DI12 form) are used by NCL staff to create MARC21 collection level records.

- The VDMS team created a Video Lab to provide a venue for transferring image data from original media to mass-storage media. It is equipped with video processing hardware and software that enables VDMS project staff to develop and document workflow processes, enable more robust file conversion and video management capabilities, and create files suitable for online access and long term retention. Final Cut Pro 5™ software is installed on two Apple™ G5 video editing workstations to allow in-house video processing by encoding the raw video data into online-accessible format.

Additional metadata is needed to assure that observations can be referenced to a specific point in the world ocean. Geographic information is often collected automatically from shipboard systems, using Global Positioning System (GPS) or other navigation technologies. The geospatial relationship of video data to positional information is maintained by using time-stamp information available from both the video and navigation sources. In addition to time-stamp annotations for individual video tapes, data managers and PIs often provide descriptive metadata for a video and/or image collection using the DV12 and/or DI12 templates. They may also provide copies of cruise reports and/or other data reports created concurrent with or subsequent to the video collection. Descriptive information about video content, technical details about file formats, encoding algorithms, and processing equipment are needed to ensure that these videos are accessible and meaningful for the long term. Related navigation data, technical details and digital reports are archived with other observation data and become accessible through the NODC Ocean Archive System (OAS) <http://www.nodc.noaa.gov/Archive/Search/>.

V. VIDEO DATA ARCHIVING AND MULTIPLE ACCESS

When video tapes and related data are received at the NCL, a metadata librarian organizes and enters the received products into online inventories and appropriate expedition folders. Figure 5 illustrates the flow of video and other data from data providers to the archive centers [3]. NCL notifies NODC that a new collection of video (and other materials) has been acquired. An NODC data content manager creates an accession entry in the NODC Accession Tracking Data Base (ATDB) [4] for the collection of tapes and related materials.

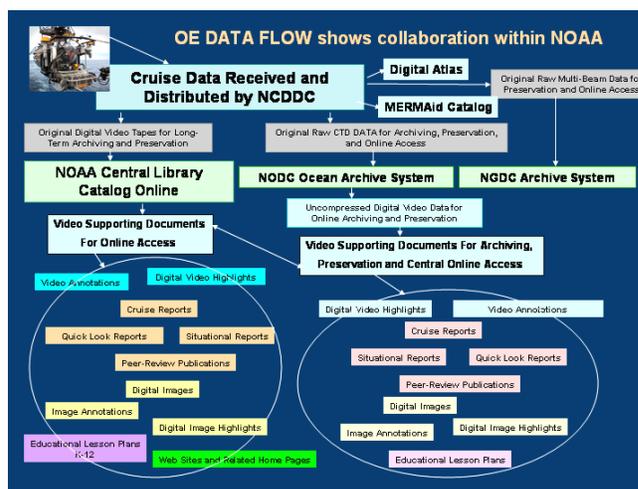


Figure 5. Schematic diagram showing high-level data flow in the NOAA VDMS (http://docs.lib.noaa.gov/OEDV/VDMS_Workflow_Ocean2005.jpg)

NODC provides long term archival storage, management and stewardship of digital oceanographic observation data and metadata. The NODC accession number assigned during the data ingest process is a tracking number for the collection. A

VI. SUMMARY

copy of clips and highlights files from specific OER cruises is added to the associated NODC accession, with a link to the file established in the NCL MARC21 record for the collection (Figure 2). Non-video observation data collected during the cruise may be included in the same NODC accession or managed in a separate accession, with a reference to the video accession.

Archival digital files maintained at NODC are stored primarily on RAID media, with offsite backups on tape media [4]. At present, original video media provided by cruise PIs or data managers are stored and maintained by the NCL. Other hard-copy information, including a binder of paper forms created during a cruise, are also stored and maintained by NODC, the NCL, or at NCDDC. As resources for digitizing these paper media become available, digital surrogates of the paper will be maintained in the NODC digital archives with related ocean observation data. Most ocean data archived at NODC can be discovered and downloaded using the NODC Ocean Archive System (OAS, online at <http://www.nodc.noaa.gov/Archive/Search/>) [4].

The OER Digital Atlas was developed by the OER/NCDDC team using Google™ Maps applications for researchers looking for deep-coral video and raw data. Users can filter the color-coded dots by year and exploration theme on the map. Each dot represents an OER signature cruise from 2001 to the present. Expedition information, including scientific observations, dive activities, ship navigational data can be tracked down and retrieved. Figure 6 illustrates the OER Digital Atlas interface.

Additional links provide access to data sets archived at the National Geophysical Data Center (NGDC), NODC or NCL, to Expedition Education Modules (EEMs) developed specifically for some NOAA signature expeditions, and to geographic information system (GIS) applications with detailed cruise and related data layers.

The VDMS project establishes a solid example of procedures to assure that NOAA's scientific video data in both physical and online formats are archived and preserved for future generations. The VDMS project working group continues to collaborate closely with NOAA OER project scientists, oceanographers, and IT specialists to develop data management requirements and strategies. This project provides an ongoing opportunity to improve the quality and completeness of metadata and information used in the NOAAALINC catalog and the NODC Ocean Archive System and to provide online access to NOAA ocean exploration video and related data to a global customer base.

The successes of the VDMS project demonstrate that much has been done, but there is more work to do. As additional resources become available, plans include providing online access to broader subsets of available digital video holdings, hosting an informal seminar series, and examining how other groups (e.g., educators, other scientists) use the digital video data. The long term VDMS Project plans include:

- Increasing access to multi-platform video images through the NOAA Libraries Online Catalog (NOAALINC) and to the WorldCat catalog. WorldCat is the world's largest and richest database of bibliographic information, linking approximately 108 million bibliographic records from the catalogs of over 54,000 libraries in 109 countries.
- Developing a web-based portal from which diverse OER ocean data, including video, still image, and audio files will be accessible via text-driven searches or from map-driven searches using a digital atlas.
- Using the NODC Archive Management System as the central digital file management repository for video, still images, ocean observations, and related documentation.
- Expanding the scope of relevant video and image data to include similar data and information from other NOAA Line Offices and Program Offices.
- Preparing the VDMS software, equipment and workflow capabilities to utilize the forthcoming high-definition video and image data from the new NOAA Ship OKEANOS EXPLORER.

VII. ABBREVIATIONS

ATDB – NODC Accession Tracking Data Base
CoRIS – NOAA's Coral Reef Information System
FGDC – Federal Geographic Data Committee
GPS – Global Positioning System
MARC21 - Machine Readable Catalog (Standards for the representation and communication of bibliographic and related information in machine-readable form.)
NCL – NOAA Central Library
NODC – National Oceanographic Data Center
NOAALINC – NOAA Library and Information Network Catalog
NCDDC – National Coastal Data Development Center
OAS – NODC Ocean Archive System

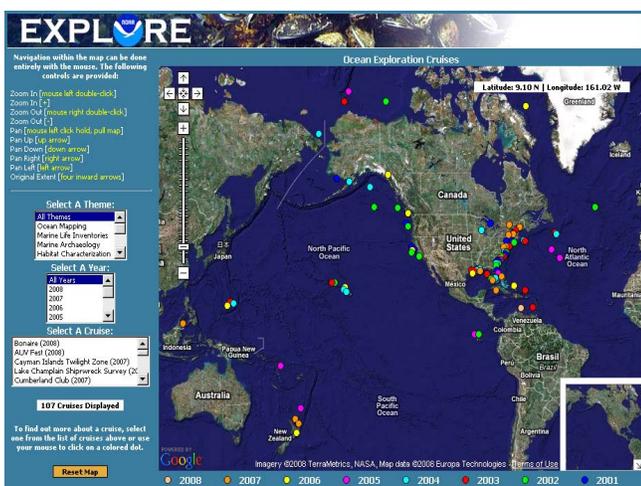


Figure 6. OE Digital Atlas using Google Maps provides the NOAA signature expeditions information

<http://explore.noaa.gov/data-management/digital-atlas>

OER – NOAA Office of Ocean Exploration and Research
OCLC – Online Computer Library Center, Inc.
PI – Principal Investigator.
MARCXML – MARC21 eXtensible Markup Language (XML
schema to represent MARC21 data elements.)
MDSD – Marine Data Stewardship Division
MERMAid – Metadata Enterprise Resource Management Aid
VDMS – Video Data Management System

REFERENCES

- [1] US Department of Commerce, National Oceanic and Atmospheric Administration, National Oceanographic Data Center, National Coastal Data Development Center, Metadata Enterprise Resource Management Aid (MERMAid), <http://www.ncddc.noaa.gov/metadatasource/metadata-tools> (Last accessed on July 14, 2008).
- [2] Reese, Terry, MarcEdit version 4.5, Free software available online at <http://oregonstate.edu/~reese/marcedit/html/>. (Last accessed on July 9, 2008).
- [3] Fiolek, A., D. W. Collins, D. Anderson, J. Beattie. 2006. NOAA Ocean Exploration Digital Video and Image Data: Archiving, Preserving, and Accessing Online Oceanographic Information. In: *Oceans 2006 MTS/IEEE, Boston: revolutionizing marine science and technology*. Columbia, Md.: Marine Technology Society; Piscataway, N.J. : IEEE. CD-ROM. Also available online http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4099116
- [4] D.W. Collins and S.B. Rutz. 2005. The NODC archive management system: archiving marine data for ocean exploration and beyond. In: *Proceedings of the Oceans 2005 MTS/IEEE Conference & Exhibition, Washington, D.C., USA: one ocean*. Columbia, Md. : Marine Technology Society ; Piscataway, N.J. : IEEE. CD-ROM. Also available online at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1640202