NOAA Satellite and Information Service
National Coastal Data Development Center (NCDDC)

Thunder Bay Sinkholes 2008
Data Management Plan

Expedition Description:
This project is focused on exploring shallow and deep water submerged sinkhole ecosystems in Thunder Bay National Marine Sanctuary in Lake Huron. The science team is planning to use a moored hydrographic instrument array and ROV and SCUBA dives for exploratory field operations.

Expedition Objectives:
This project is expected to provide a clear working picture of the bathymetry, physico-chemical conditions, hydrologic processes, variety and distribution of life in a range of submerged sinkhole ecosystems in this Laurentian Great Lakes Basin. The detailed objectives are to:

1. Map the underwater habitats using ROV and SCUBA divers (sensors and cameras),
2. Evaluate the variable physico-chemical conditions for one year (continuous time-series exploration),
3. Measure the rate of groundwater discharge (Ra, Rn tracers and stable oxygen isotopes),
4. Quantify ground water and sediment pore water nutrient chemistry (colorimetry and ion chromatography),
5. Identify the carbon composition of benthic mats and sediments (stable carbon isotopes in sediment cores),
6. Describe the fine-scale vertical organization of the microbial mats (epifluorescence microscopy, molecular characterization),
7. Identify the metazoan sinkhole inhabitants (light microscopy and molecular characterization),
8. Study mat growth (light/dark growth experiments).

Expedition Principals for Data Management:
Steve Ruberg, NOAA/GLERL (Principal Investigator)
Anne Clites, NOAA/GLERL
Susan Gottfried, NCDDC

Anticipated Data Submission for Archive:

- Quick Look Report (form to be supplied)
- Vessel navigation and hull-mounted sensor data in non-proprietary format (ASCII text)
- Submersible navigation and integrated sensor data in non-proprietary format (ASCII text)
- Copies of video collected during dive operations (e.g. DVD, MiniDV, etc)
- Highlight image CD (15-20 still images with captions/credits)
- Highlight video clip
- Mapping raw data sets
- CTD Cast raw data (ASCII text)
- Sample logs
- Dive logs (ROV Watch logs, SCUBA diver log)
- Cruise Summary report
- Publications, Journal Articles, etc.

Data Management Plan:

Thunder Bay Sinkholes 2008 is considered a Signature Expedition by the Office of Ocean Exploration & Research (OER). As such, and according to the OER Data Management Guidance (Appendix A), OER is providing an Ocean Exploration (OE) Data Manager for the expedition. The OE Data Manager will use the Cruise Information Management System (CIMS) software to document the metadata, the who, what, when, where, why and how of any raw data sets, samples, multimedia files, and products collected, recorded, and/or produced during and after the Thunder Bay Sinkholes 2008 expedition. The OE Data Manager will enter information into CIMS, which was developed specifically for documenting the science activities onboard OE expeditions and producing Federal Geographic Data Committee (FGDC) compliant metadata records. OER is mandated by the government to archive data from its expeditions at the appropriate NOAA National Data Center within a reasonable amount of time; FGDC-compliant metadata is required to be included with the corresponding data when it is submitted to be archived. CIMS will generate the metadata within the Metadata Enterprise Resource Management Aid (MERMAid) tool maintained at the NOAA National Coastal Data Development Center (NCDDC). Once in MERMAid, the metadata will be ready to be published for public discovery and access at such time that is appropriate. The metadata generated through CIMS will have the background of the mission, its participants, its objectives and other information, and thus will provide more robust and more complete referential information for the consumer.

Several types of metadata records are generated for each expedition. The highest-level record is called a Collection-level metadata record and contains the most basic header information about the expedition – title, vessel, chief scientist, mission abstract, dates, geographic region, and keywords or phrases that enhance discovery of the metadata. More detailed metadata records document the raw data sets recorded, samples collected, multimedia images or video recorded, and products generated as a result of the expedition. The collection-level metadata record enables the archive centers to prepare their systems to receive associated data sets that are their responsibility to archive. The National Oceanographic Data Center archives oceanographic, biological, chemical, and environmental data sets. The National Geophysical Data Center archives geological and geophysical data sets. The NOAA Central Library archives physical multimedia, digital multimedia files (such as video highlights and highlight images), and paper and digital products (such as reports, journal articles, publications, etc). See Figure 1 below for an illustration of the flow of data from CIMS to the archive centers.

Once published in MERMAid, metadata records are harvested by Geospatial One-Stop, the government's preferred repository for geospatial metadata. MERMAid is also considered a node on the National Spatial Data Infrastructure (NSDI) clearinghouse and metadata records published from MERMAid will be distributed across this network for increased discoverability through metadata search capabilities.

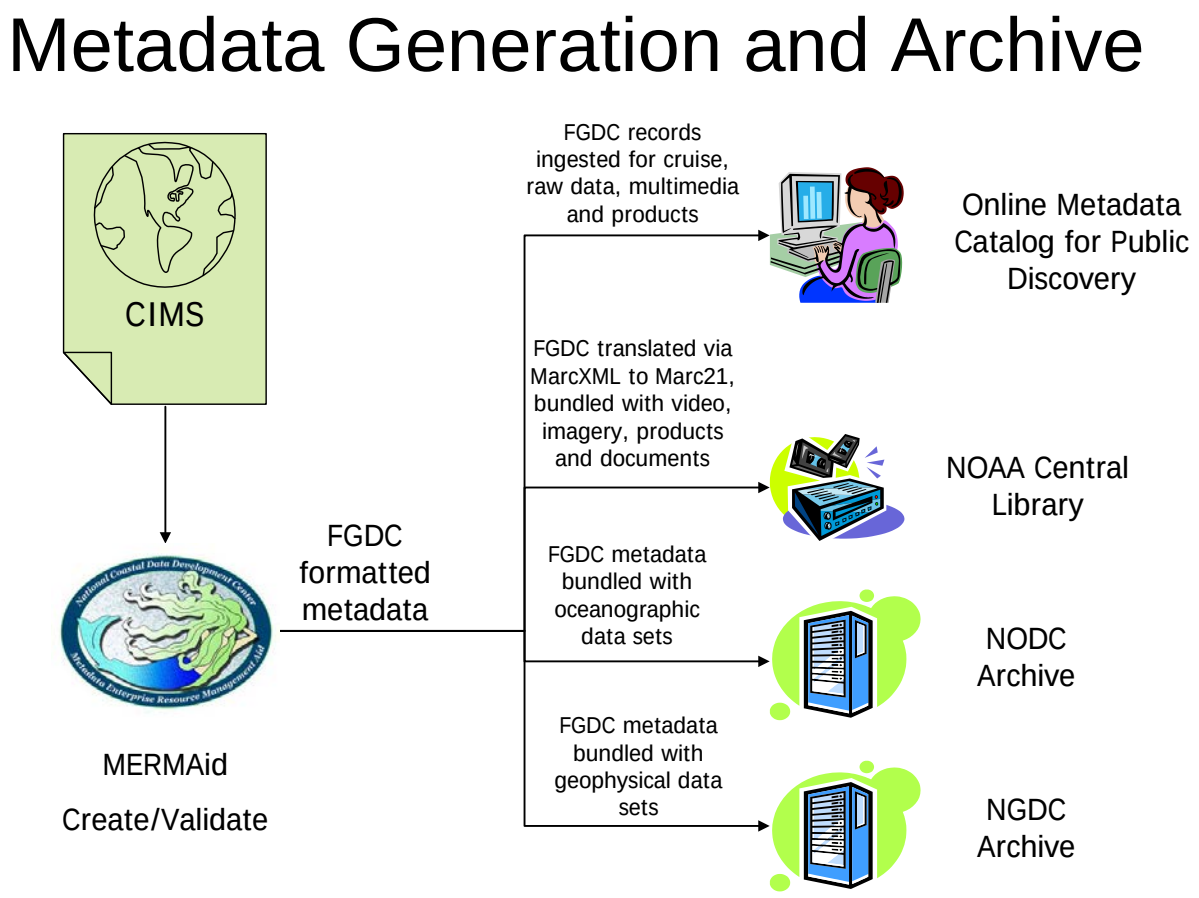

Fig. 1: Metadata Flow of CIMS for Archive, Discovery and Access

Thunder Bay Sinkholes 2008 Metadata Hierarchy

For the Thunder Bay expedition, the OE Data Manager will populate an overarching metadata record called a Cruise record with the abstract of the entire mission, the names of the chief scientist and other participating personnel, and the temporal and spatial boundaries. There are certain cases when there are multiple Projects related to a Cruise. In the case of the Thunder Bay Sinkholes expedition, each research site, Isolated Sinkhole, Middle Island Sinkhole, and Misery Bay, has unique objectives and unique temporal and spatial boundaries and it is logical to separate the three research sites into unique Projects within CIMS. Project metadata records will be created for each. Project metadata records are populated with the information provided by the research team through the proposal process - the names of the principal investigators, the project objectives, the instrumentation and sensors being employed, and the target research sites. See Figure 2 below for the metadata hierarchy.

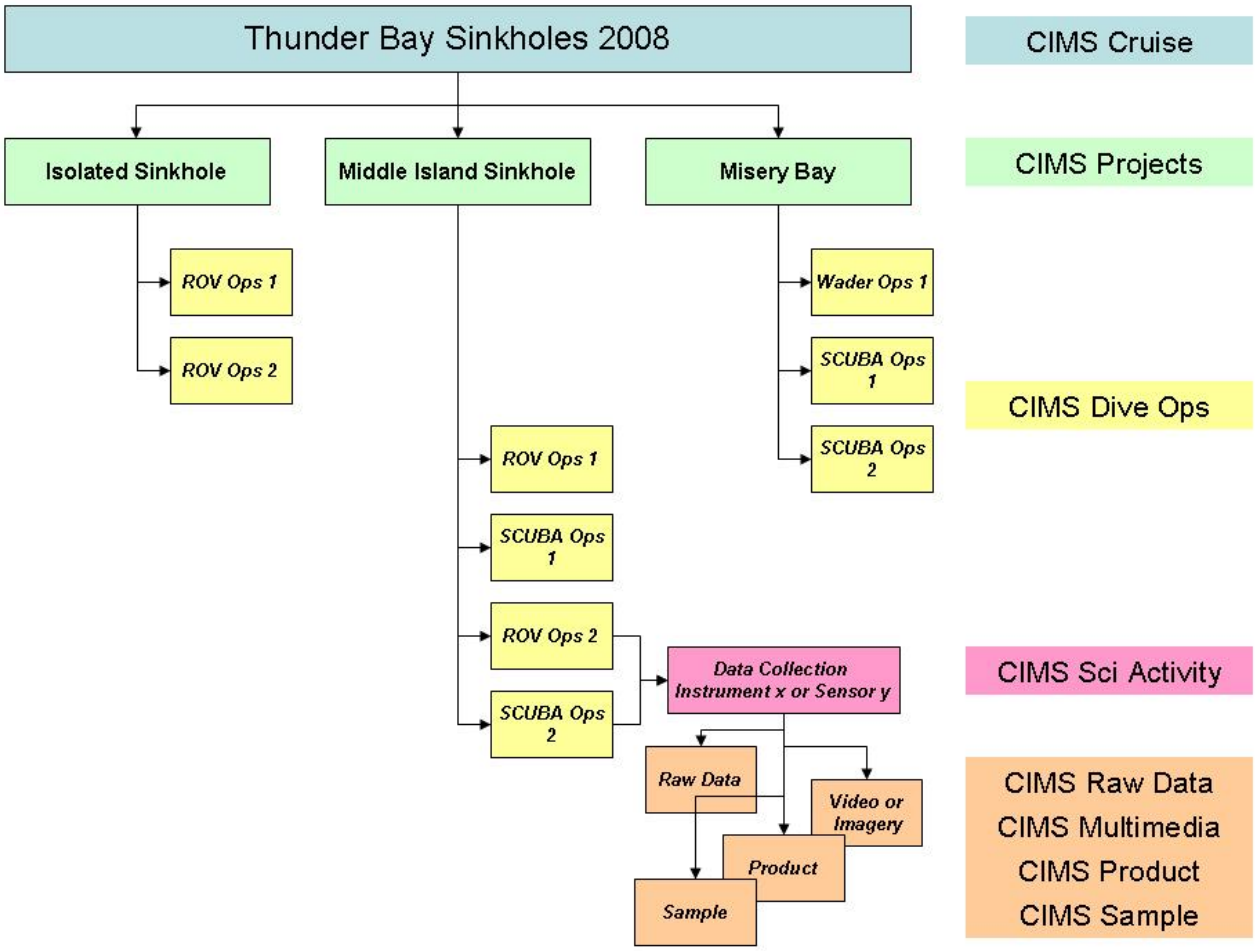

Fig. 2: Metadata Hierarchy within CIMS for Thunder Bay Sinkholes

The OE Data Manager will document the metadata from dive operations, whether ROV or SCUBA, and the science activities that occur either during dive operations or independently (e.g. CTD casts, hull-mounted sensors, net tows). Science activities are defined as the use of some instrument or sensor to record or collect data - samples, multimedia files, or raw data sets. As an example, if the ROV's onboard color video camera records video on DVD, the OE Data Manager will create a multimedia metadata record within CIMS and associate it with the science activity metadata record detailing the use of the video camera, which will be associated with a dive operation metadata record for the ROV ops. Similarly, a non-dive related science activity might be the deployment of a CTD cast. In this case, the OE Data Manager will create a raw data metadata record and associate it with the science activity metadata record detailing the use of the CTD device.

Within CIMS, metadata records can also be created for shuttling operations, vessel transits, media events, education/outreach events, incidents or accidents, weather events, marine mammal sightings, unscheduled down times, and others. Many of these are useful for the daily Situation Reports (SitRep), an operational document that typically gives the current position of the vessel, the personnel onboard, the operations, activities and events that occurred during the reporting period and a first-person summary of the expedition during the reporting period. SitReps are considered operational

documents and are emailed to the OE Expedition Coordinator as well as to any colleague or family member, whose email address has been supplied to the Data Manager.

Geospatial and other Products for Discovery

NCDDC maintains a Digital Atlas for OER that provides a geographical user interface for data discovery and access to OE-sponsored expeditions. The tool, which can be accessed online at http://www.ncddc.noaa.gov/website/google_maps/mapsOE.htm, color codes the expeditions by year and provides filters by year and by exploration theme to affect the display. Clicking on an expedition icon will display a bubble of information. The first screen is a brief summary of the expedition and a link to its web page on http://oceanexplorer.noaa.gov, OER's Education and Outreach website. Once the data from the expedition is archived, a 'Links' tab will be available with instant access by the user to the location of the expedition's data at each of the archive centers. If an Expedition Education Module (EEM) was created by OER's Education Program for the expedition, a link to the EEM will also be available on this 'Links' tab. An 'Atlas' tab will be available if the expedition has been included in a more sophisticated Geographic Information System (GIS) application with an Interactive Map Service allowing the user to view information and data from the expedition in relation with other data sources in the same study area. Navigational and underway data from the vessels and submersibles provide the initialization data necessary to be included in the GIS. Once incorporated, other data sets or products can be accessible to the user.

The NOAA Central Library maintains a Video Data Management System (VDMS) of all physical media archived in their climate-controlled storage rooms. Each media unit is catalogued in the library and is accessible for viewing, with the principal investigator's permission, in the Library's Media Room.

The NODC assigns Accession numbers to the data sets they receive for their Ocean Archive System (OAS). Visitors to the NODC website are able to discover the data sets archived at NODC through the OAS user interface, accessible through http://www.nodc.noaa.gov.

The NGDC uses the GEOphysical DAta System (GEODAS), http://www.ngdc.noaa.gov/mgg/ to archive and provide access to multibeam bathymetry and other geophysical and geological data.

Appendix A: Data Management Guidance

NOAA's Office of Ocean Exploration & Research (OER) is in accord with the 2000 President's Panel on Ocean Exploration call for a Federal program that promotes data management and dissemination to ensure that discoveries can have maximum impacts in the research, commercial, regulatory, and educational realms. OER advocates open scientific communication and expects significant findings from supported research and educational activities to be promptly submitted for publication with authorship that accurately reflects the contributions of those involved. It expects PIs to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work. It also encourages grantees to share software and inventions, once appropriate protection for them has been secured, and otherwise act to make the innovations they embody widely useful and usable. OER will put these principles into practice, in ways appropriate to field and circumstances, through the proposal review process; through award negotiations and conditions; and through appropriate support and incentives for data cleanup, documentation, dissemination, storage and the like. Adjustments and, where essential, exceptions may be allowed to safeguard the rights of individuals and subjects, the validity of results and the integrity of collections, or to accommodate legitimate interests of investigators.

Each applicant should articulate his/her willingness to comply with these principles. The proposal Project Description should outline the plans for preservation, documentation, and sharing of data, samples, physical collections, curricula materials and other related research and education products. Plans for handling of data and other products will be considered in the review process.

Grant recipients are responsible for archiving all acquired data sets and associated products in the appropriate NOAA archive (i.e., National Oceanographic Data Center, National Geophysical Data Center, National Coastal Data Development Center, National Climate Data Center, and the NOAA Central Library) as soon as practical and, in no case, later than two years following the completion of the expedition. Within 60 days of completion of the expedition/field work/etc, grant recipients should provide a collection level metadata record (e.g., number & type of data, and description of the data collected) to the OE Data Manager for submission to the National Data Clearinghouse. This record should be created in compliance with the Federal Geographic Data Committee Content Standards for Digital Geospatial Metadata (FGDC-STD-001-1998) in accordance with Executive Order 12906, Co-ordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure dated April 1994 http://www.fgdc.gov/policyandplanning/executive_order.

For continuing observations, data inventories should be submitted periodically if there is a significant change in location, type or frequency of such observations. Archiving of any specific archaeological site information that is considered sensitive material under Section 304 of the National Historic Preservation Act of 1966 (as amended) will not be required to be in compliance with these requirements.

For a subset of funded expeditions, OE will facilitate data collection, management and archival by providing the in-field services of an OE Data Manager. The OE Data Manager will be responsible for documenting all expedition activities (e.g., science, education and outreach, transits, shuttles) in the OE Cruise Information Management System (CIMS). The metadata records contained in the CIMS will be provided to the Science Team for inclusion in required reports, and will be distributed to the appropriate NOAA archive.

To aid in the Census of Marine Life (CoML; http://www.coml.org) goal of determining the distribution, abundance and diversity of marine life, all projects are encouraged to maintain their biological species information on-line and linked (either directly or via a Federal data center) to the CoML database, the Ocean Biogeographic Information System (OBIS) (http://www.iobis.org). A statement to this effect should be included in the proposal. Biological samples should be registered with the Census of Marine Life's OBIS database (http://www.iobis.org) or a Federal database that is linked to OBIS.

Other data or data products may also be requested at the discretion of the OE Director for administrative, scientific, or educational use. All such product requests will be considered in full collaboration with the applicant, and any data sensitivity issues will be handled appropriately. PIs and their institutions are responsible for meeting all legal requirements for submission of data and research results that are imposed by foreign governments as a condition of that government's granting research clearances. Each PI and institution must determine their legal obligations in this respect, with the assistance of the Department of State and NOAA, as necessary.

The applicant is encouraged to consider budgeting appropriate resources for any anticipated costs to comply with these data management principles. This generally includes budgeting for the following types of activities: reports, reprints, page charges or other journal costs (except costs for prior or early publication); necessary illustrations; cleanup, documentation, storage and indexing of data and databases; development, documentation and debugging of software; and storage, preservation, documentation, indexing, etc., of physical specimens, collections or fabricated items. The OE Data Manager, when not otherwise specifically designated in each Grant or Contract award, will be the NOAA National Coastal Data Development Center's OER Liaison Officer (228-688-2936 or toll free 1-866-732-2382).