Incorporation of Gulf of Mexico Benthic Survey Data into the Ocean Biogeographic Information System

Robert S. Carney
Department of Oceanography & Coastal Sciences
Louisiana State University

MMS Securing Ocean Energy & Economic Value for America

LSU

Coastal Marine Institute

History of Benthic Data Archiving

- An Input – Input System
  - NODC archives of benthic data virtually unused
  - Mode of archiving did not anticipate any actual use of the data
  - No tools provided for data access
  - Very different for physical oceanography data which has been archived and used to create very useful data products

The Future of Benthic Data Archiving

- A Wealth of Basic and Management Questions
  - What commonalities exist in the distribution of ocean species from the local to global scale?
  - What commonalities exist on the distribution of biodiversity over the same scales?
  - What is the most useful classification of exploited benthic systems?
  - What types of long-term change are occurring?
  - Can better surveying and monitoring be devised?

The Question Answering Tools

- BIOINFORMATICS
  - A rapidly growing suite of concepts and software specifically designed for examination of existing data (data mining)
  - Made very practical by:
    - Fast computers capable of addressing very large amounts of data
    - Database interconnection via the Internet
    - Extremely flexible data format requirements
  - Several local and international efforts of which Ocean Biogeographic Information Systems (OBIS) is an important one

Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.iobis.org/ Go OBIS rutgers

Getting Started | Latest Headlines | The Siboga Expedition | Amethyst Cabochon... | Welcome to Procure... | MEB: Wiebke Ziebis | Louisiana State Univ...

NODC Taxonomic Code | http://www.iobis.org/

OBIS
OCEAN BIOGEOGRAPHIC INFORMATION SYSTEM

How to... | Resources | News | About Us

EXPLORE DATA ON LOCATIONS OF MARINE ANIMALS AND PLANTS

Obtain data tables, maps and predict distributions using environmental information

10.3 million records of 75000 species from 153 databases

SEARCH BY NAME

"Great white shark" or "Carcharodon" or "Carcharodon carcharias"

Search >>

Advanced Search including date, depth, dataset Browse by taxonomic groups

SEARCH BY GEOGRAPHY

Click on the map to set the query box 5° search area Search >>

Transferring data from mt3.google.com...

Ocean Biogeographic Information System - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.iobis.org/OBISWEB/ObisControllerServlet?searchName=Paraprionospio+pinnata&category=all&name

Go OBIS rutgers

Getting Started | Latest Headlines | The Siboga Expedition | Amethyst Cabochon... | Welcome to Procure... | MEB: Wiebke Ziebis | Louisiana State Univ...

NODC Taxonomic Code | Ocean Biogeographic Information Sy...

OBIS OCEAN BIOGEOGRAPHIC INFORMATION SYSTEM

How to... | Resources | News | About Us

Name Search | Search >>

Advanced Search | Browse by taxonomic groups

Would you like to give us any feedback on OBIS? Click here to take our survey.

RESULTS

How to cite OBIS data

Sort by: Scientific name | Common name | Higher taxon | Number of global records

Search criteria included:
" Paraprionospio pinnata "

Results 1 - 1 (of 1)

Paraprionospio pinnata
a polychaete, name verified, 1180 global records

Result page: **1**

OBIS is a project of the CENSUS OF MARINE LIFE

OBIS strives to document the oceans' diversity, distribution, and abundance of life. © Rutgers, The State University of New Jersey, 2002-2006.

Done

File Edit View Go Bookmarks Tools Help

http://www.iobis.org/OBISWEB/ObisControllerServlet?searchCategory=/AdvancedSearchServlet&genus=Parapri

Go OBIS rutgers

Getting Started Latest Headlines The Siboga Expedition Amethyst Cabochon... Welcome to Procure... MEB: Wiebke Ziebis Louisiana State Univ...

NODC Taxonomic Code | Ocean Biogeographic Information Sy...

SPECIES INFORMATION

How to cite OBIS data

Back to previous page

Notes to users

Paraprionospio pinnata

Name verified: Catalogue of Life; ITIS

Organism type: a polychaete ITIS

***Paraprionospio pinnata* Data Extent Map** (from OBIS Australia/ C Square Mapper)

For more options, view full-size map
or use other mapping systems:

KGSMapper: a dynamic habitat analysis and inference tool from Biogeoinformatics of Hexacorals

(if your request contains < 1,000 records, taxon names will be transmitted with the locations; if > 1,000 records, only locations will be sent)

KGS »

Additional Information on *Paraprionospio pinnata*

- Barcode of Life
- GenBank
- Catalogue of Life
- ITIS
- Google Images
- Google Scholar
- uBio

Done

Nature of Benthic Survey Data

- Species
  - Full taxonomic position
- Where Found?
  - Position
  - Depth
  - In relationship to sampling design
  - How sampled
- How Many?

OBIS Data Elements

Based on the Darwin Schema

Date Last Modified
Institution Code
Collection Code
Catalog Number
Record URL
Scientific Name
Basis of Record
Source
Citation
Kingdom
Phylum
Class
Order
Family
Genus
Subgenus
Species
Subspecies
Name
Author
Identified By
Year Identified
Month Identified
Day Identified
Type Status
Collector's Number

Field Number
Collector Year
Collected Month
Collected Day
Collected Start Year
Collected Start Month
Collected Start Day
Collected End Year
Collected End Month
Collected End Day
Collected Julian Day
Start Julian Day
End Julian Day
Time of Day
Start Time of Day
End Time of Day
Time Zone
Continent or Ocean
Country State or Province
County Locality
Longitude
Start Longitude
End Longitude
Longitude
Start Longitude
End Longitude

Longitude
Start Longitude
End Longitude
Latitude
Start Latitude
End Latitude
Coordinate Precision
Start/End Precision
Bounding Box
Minimum Elevation
Maximum Elevation
Minimum Depth
Maximum Depth
Depth Range
Temperature
Sex
Life Stage
Preparation
Type
Individual Count
Observed Count
Observed Weight
Previous Catalog Number
Relationship Type
Related Catalog Item
Notes

MULDARS F-132

- **MUL**ti-**D**iscipline **A**rchives **R**etrieval **S**ystem
  - NODC required format 1970 – 1995
  - Strong emphasis upon compact data storage
  - Little emphasis upon data retrieval
  - Designed during era of punched card input
    - Only 80 characters per card
  - F-132 required format for benthic fauna

ata Formats and Codes - Mozilla Firefox
Edit View Go Bookmarks Tools Help
http://www.nodc.noaa.gov/General/NODC-datafmts.html
Go muldars code
Getting Started Latest Headlines The Siboga Expedition Amethyst Cabochon... Welcome to Procure... MEB: Wiebke Ziebis Louisiana State Univ...

NOAA Satellite and Information Service

tional Oceanographic Data Center

>> Search NODC
>> Contact NODC

u are here: NODC Home > Access Data > Data Formats & Codes

Ocean Data

- Access Data
- CD-ROMs & DVDs
- Publications
- NODC Data by organizational component
- NOAA Library Resources
- **Data Formats & Codes** details at right >>
- Online Store
- Ordering & payment procedures

>> NODC Formats & Codes

The following is a table of the principal NODC data storage (or data output) formats and codes. This list covers only data stored in the NODC archive databases; it does not include formats for all NODC data collections on CD-ROM or other online projects. CD-ROM data sets and online projects are described in documentation files included with each data product.

- NODC Platform (Ship) Code List
- List of codes used in the World Ocean Database - includes Institute codes, Country codes, Ship codes, etc...
- Complete NODC File Type Code List (all codes in the table below listed in one file).
- Taxonomic Code for Biological Species - this link will send you to ITIS, as NODC no longer maintains this code. We also have information about the old Taxonomic Code CD-ROM that is now out-of-date.

DATABASE / DATA FILE NAME	RECORD LAYOUT	CODE TABLES
Oceanographic Profile Database		
Oceanographic Profile Data Format	P3	
Station Data II Output Format	SD2	SD2 Code Tables
Universal Bathythermograph Output Format	UBT	UBT Code Tables
High-Resolution CTD/STD Output Format	F022	F022 Code Tables
World Ocean Database	WOD	

start Adobe Illustrator 3 Microsoft Office ... Data Formats and Co... Microsoft PowerPoint...

Data Formats and Codes - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.nodc.noaa.gov/General/NODC-datafmts.html
Go muldars code
Getting Started Latest Headlines The Siboga Expedition Amethyst Cabochon... Welcome to Procure... MEB: Wiebke Ziebis Louisiana State Univ...

Sea Level Data		
Hourly	F184	F184 Code Tables
Daily	F185	F185 Code Tables
Monthly	F186	F186 Code Tables
Surface Current Data System	SCUDS	NONE
Principal MULDARS Data Formats*		
Current Meter Data (Resultants)	F005	NONE
Current Meter Data (Components)	F015	F015 Code Tables
Ocean Drifter Data	F156	F156 Code Tables
Water Physics and Chemistry	F004	F004 Code Tables
Marine Chemistry	F069	NONE
Marine Toxic Substances and Pollutants	F144	F144 Code Tables
Phytoplankton	F028	Taxonomic Code
Zooplankton	F124	F124 Code Tables
Primary Productivity 1	F029	F029 Code Tables
Primary Productivity 2	F049	F049 Code Tables
Inter-tidal Organisms and Habitats	F030	F030 Code Tables
Inter-tidal/Sub-tidal Organisms and Habitats	F100	F100 Code Tables
Benthic Organisms	F132	F132 Code Tables
Fish/Shell-fish Surveys	F123	F123 Code Tables
Marine Mammal Sighting and Census	F127	F127 Code Tables
Wind Measurements from Buoys	F101	F101 Code Tables
Pressure Gauge Data	F017	F017 Code Tables

(*Data of these types received after 1994 are stored in originators' formats. See the NODC Ocean Archive System for data stored exactly as we received it.)

http://www.nodc.noaa.gov/General/NODC-Archive/f030.txt

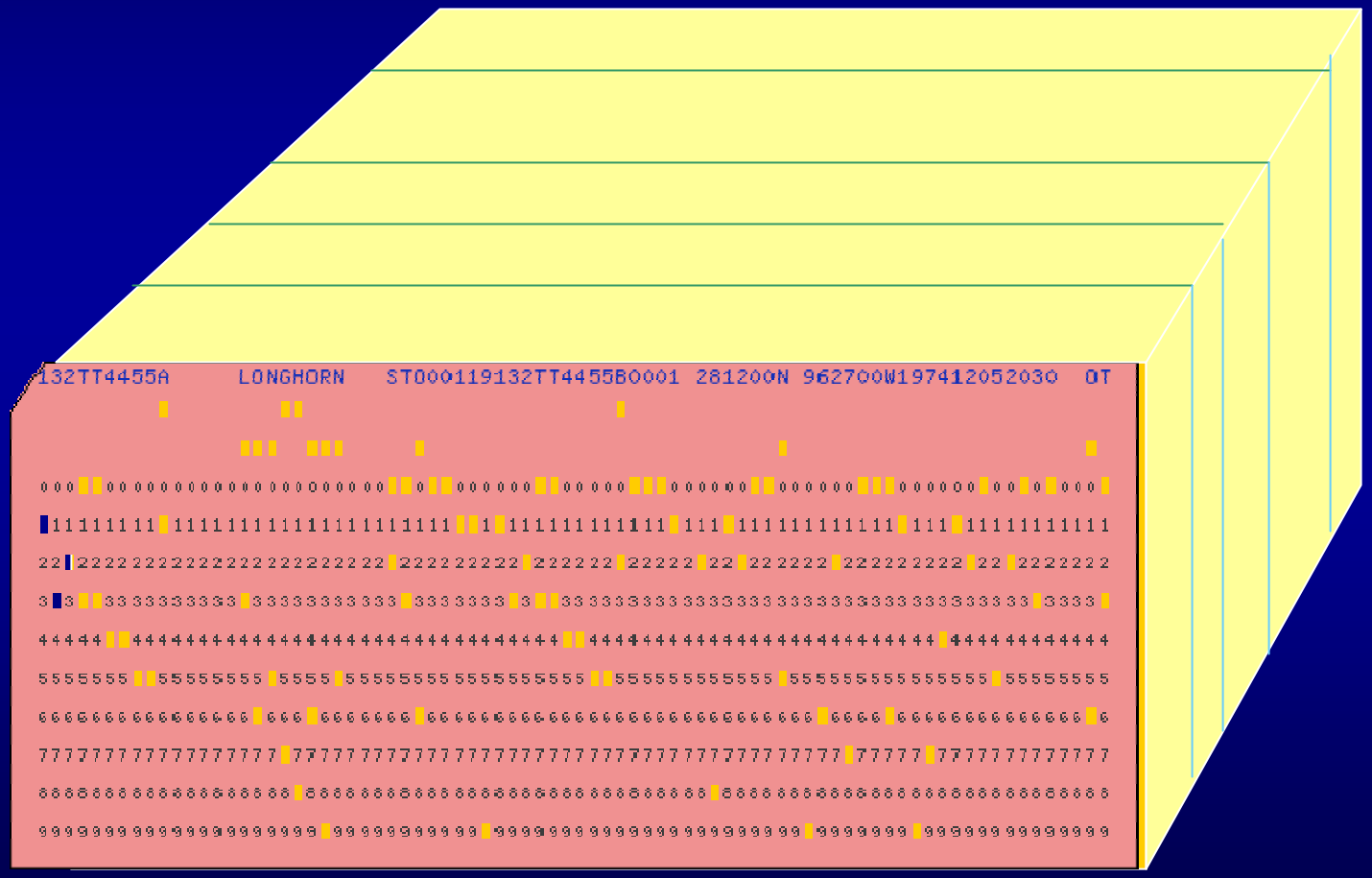
MULDARS Files Use Multiple Headers

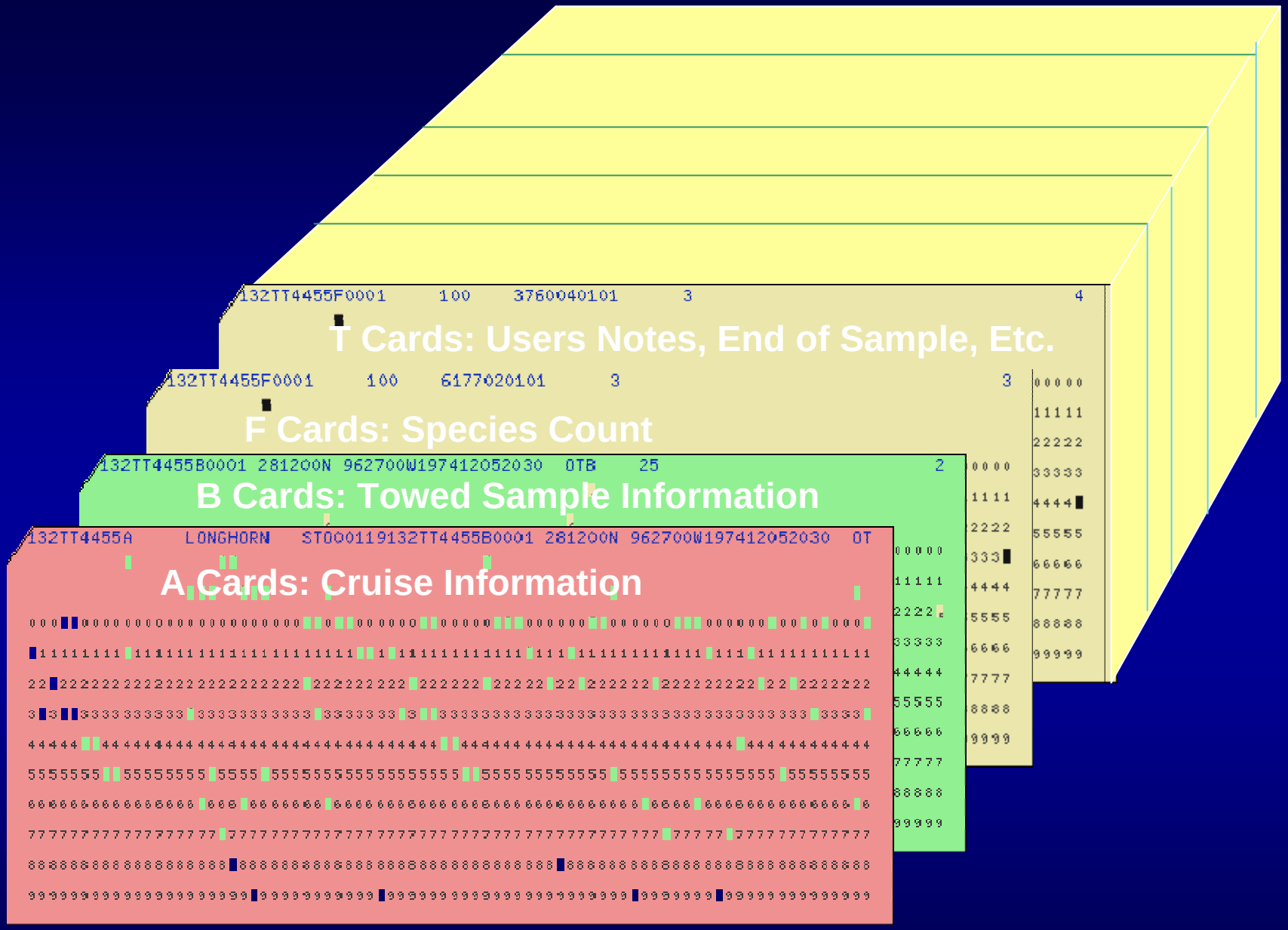
- HEADER Type 1 (File Information)
  - HEADER Type 2 (Sample Information)
    - DATA (Specific Information)
    - DATA (Specific Information)
    - DATA (Specific Information)
  - HEADER Type 2 (Sample Information)
    - DATA (Specific Information)
    - DATA (Specific Information)
    - DATA (Specific Information)

F-132 3-Header Content

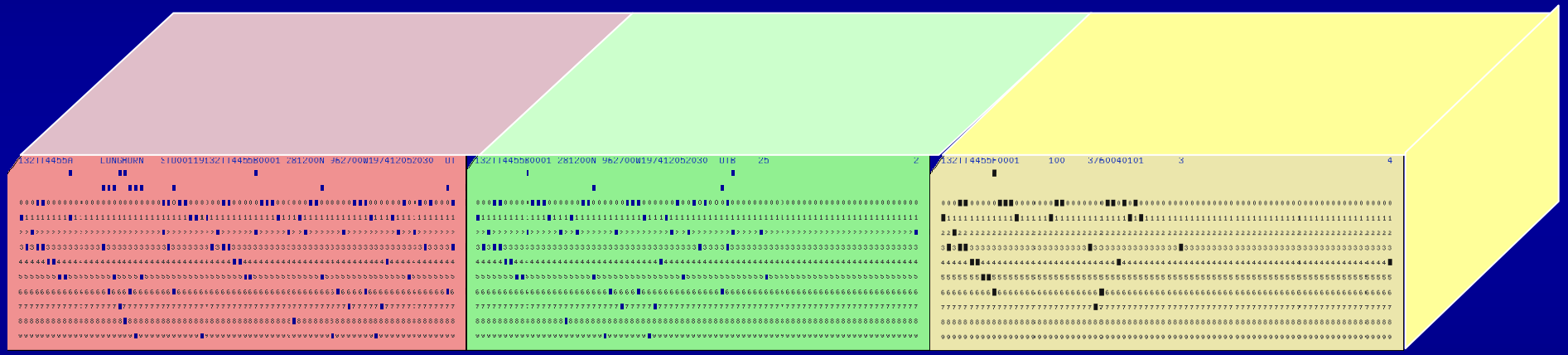
- A – records (1 per file)
  - Ship, Cruise, Project, Investigator, etc.
- B – records (1 per sample)
  - Towed sample, date, lat., long., depth, device, etc.
- C – records (1 per sample)
  - Point sample, date, lat, long, device, etc.
- F – records (all taxa in sample)
  - Species (coded), count, weight
- T – records (1 per sample or as needed)
  - Text, originator's notes

Data Constrained by Punched Cards




132TT4455F0001 100 3760040101 3 4
T Cards: Users Notes, End of Sample, Etc.
132TT4455F0001 100 6177020101 3 3
F Cards: Species Count
132TT4455B0001 281200N 962700W197412052030 OTB 25 2
B Cards: Towed Sample Information
132TT4455A LONGHORN ST000119132TT4455B0001 281200N 962700W197412052030 OT
A Cards: Cruise Information

A Storage Savings of 2/3rds



Basic Task of Data Transfer

1. Obtain F-132 files from NODC (online)
   - Verify Content with Excel
2. Concatenate small files into large
   - Windows Command Prompt
3. Add cruise (A record) and station (B/C records) information to each Species (F record) count
   - Requires Custom Programming
4. Convert NODC Taxonomic Code to species name and full taxonomic code
   - Requires Custom Programming
5. Compose OBIS Format
   - Excel

Taxonomic Information

- Species-level identification still recognized as required for characterization and monitoring
- Necessary for estimation of bio-diversity
- Necessary for new taxonomic diversity indices
- Must be accurate and consistent for archived data to be useful
- Link to data on species ecology

Breadth of Taxonomic Data

- Name = Genus + species
- Hierarchy = Genus → Kingdom
- Author
- References
- Museum Specimens
- Identification Aids
- **TOO MUCH FOR SINGLE DATA BASE**

Taxonomic Code Link

- NODC Taxonomic Code
  - First published in 1977
  - Last updated in 1996
- ITIS - Integrated Taxonomic Information System
  - USDA based
  - Participants in global inventories
  - Does not support easy conversion between code systems

NODC Taxonomic Code - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.nodc.noaa.gov/General/CDR-detdesc/taxonomic-v8.html
Go nodc taxonomic
Getting Started Latest Headlines The Siboga Expedition Amethyst Cabochon... Welcome to Procure... MEB: Wiebke Ziebis Louisiana State Univ...

Introduction to Taxonomic Code Ver. 8.0 on CD-ROM

Note: This CD-ROM is no longer available. Please see the ITIS link below.

Up to 1996, while it was in active use, the NODC Taxonomic Code was the largest, most flexible, and widely used of the various coding schemes which adapted the Linnean system of biological nomenclature to modern methods of data storage and retrieval. It was based on a system of code numbers that reflected taxonomic relationships. Over the years the problems and shortcomings created by this approach became more apparent. At the same time, relational database management systems were developed that enabled efficient searches and retrievals to be conducted without the need for taxonomic information to be built into the code. Therefore, in 1996, NODC released the final version (#8) of the NODC Taxonomic Code on CD-ROM. This CD-ROM, **now out of date**, provided the old NODC codes along with their new ITIS Taxonomic Serial Numbers to facilitate the transition to a new Integrated Taxonomic Information System (ITIS). From that point in time the NODC code was frozen, and ITIS assumed responsibility for assigning new TSN codes, and verifying accepted scientific names and synonyms. ITIS is supported by a partnership of U.S. and international organizations, including NOAA/NODC. ITIS contains nearly 300,000 entries for terrestrial, marine, and freshwater species from all biological kingdoms. While it focuses on North American species (driven by the priorities of the current ITIS partners), it also includes worldwide treatment of selected groups of fishes, birds, reptiles, molluscs, corals, and other groups. ITIS is continually seeking new partnerships, and expanding its coverage.

History of the NODC Taxonomic Code

In 1972 Richard Swartz, Marvin Wass, and Donald Boesch published A Taxonomic Code for the Biota of Chesapeake Bay at the Virginia Institute of Marine Science (VIMS Special Scientific Report No. 62). Their efforts were specifically oriented toward development of a universally acceptable coding system since, as they succinctly put it, they felt that "everyone [should] use the same code." The VIMS codes contained a maximum of 10 digits, with each

Done

Integrated Taxonomic Information System - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.itis.usda.gov/ Go itis taxonomic

Getting Started | Latest Headlines | The Siboga Expedition | Amethyst Cabochon... | Welcome to Procure... | MEB: Wiebke Ziebis | Louisiana State Univ...

ITIS Integrated Taxonomic Information System

- What's New
- About ITIS
- Data Access
- Submit Data
- Tools
- TRED
- Links
- Comments

NEW ITIS now has a new URL! Please use **http://www.itis.gov** from now on!

Welcome to ITIS, the Integrated Taxonomic Information System! Here you will find authoritative taxonomic information on plants, animals, fungi, and microbes of North America and the world. We are a partnership of U.S., Canadian, and Mexican agencies (ITIS-North America); other organizations; and taxonomic specialists. ITIS is also a partner of Species 2000 and the Global Biodiversity Information Facility (GBIF).

Quick search on:

(•) Any Name or TSN* ( ) Common Name ( ) Scientific Name ( ) TSN*

In: every Kingdom containing [ ] Search

* Taxonomic Serial Number (TSN)

Go to Advanced Search and Report

Last Updated: Sunday, 19-Nov-2006 15:23:20 MST

Privacy statement and disclaimers

http://www.itis.usda.gov/index.html

Integrated Taxonomi... | Adobe Illustrator | Microsoft PowerPoint ... | 1:24 PM

NODC "Intelligent" Code

- **5001431701**
- **50** Annelida Phylum
- **5001** Polychaeta Class
- **500143** Spionidae Family
- **50014317** *Paraprionospio* Genus
- **5001431701** *P. pinnata* Species
- Code contains taxonomic hierarchy with 10 numbers (12 for subspecies)

ITIS “Dumb Code”

- *Paraprionospio pinnata* = **:66937**
  - Limited to six alphanumeric characters
  - Code assigned as species entered regardless of taxonomic position
  - No information about species in the code
  - Potential for unlimited information in the IT IS database accessed via this code

The Transfer Experience

- Pequegnat Deep Sampling
  - No digital archive, hand entered from hardcopy
- South Texas Baseline Study
  - Excellent archive files
- South West Florida Benthic Study
  - Excellent archive files but very hard to find
- Northern Gulf of Mexico Continental Shelf Study
  - NODC files incomplete, LGL data files provided

Recommendations

1. MMS should continue policy of requiring data archiving and submission on voucher specimens.
2. MMS should require archiving in OBIS format to assure future utility.
3. Conversion is a very time-consuming task, indicating that priorities must be set.
   - Highest management priority should go to areas where long-term studies may be undertaken.
   - Highest science priority should go to areas where data are otherwise few.