DEEPEND: Deep Pelagic Nekton Dynamics of the Gulf of Mexico

Data Management Plan

Section 1: Research Consortium Information

1. Project Title:

DEEPEND: Deep Pelagic Nekton Dynamics of the Gulf of Mexico

2. Lead Principal Investigator and Contact Information:

Dr. Tracey Sutton
Nova Southeastern University
8000 N Ocean Drive
Dania Beach, FL, 33004
(954)262-3692
Tsutton1@nova.edu

3. Data Manager and Contact Information

Dr. Matthew Johnston
Nova Southeastern University
8000 N Ocean Drive
Dania Beach, FL, 33004
(954)262-3641
johnmatt@nova.edu

4. Data Manager Roles and Responsibilities

Dr. Matthew Johnston will make decisions regarding data and web management in collaboration with Dr. Tracey Sutton, lead PI and Director of the DEEPEND Consortium, and

consortiums researchers as needed. Dr. Johnston's position as data manager is an 80% full-time equivalency (FTE) appointment per year over the three year duration of the project. Individual PIs will be responsible for the submission of DIFs and data, and also the creation of metadata, overseen by the data manager.

5. Data Integrity

The integrity of dataset packages submitted to GRIIDC will be the responsibility of each of the PIs for their respective dataset. The data manager will provide guidance and oversight.

6. Data Infrastructure

Raw Data Storage:

The majority of the consortium raw and some in-process data (see 'other data' section for exceptions) will be housed on the 'NSU Works' Content Management System, hosted by Bepress via the Nova Southeastern University library spanning the duration of this project.

About NSU Works

Storage Capacity: NSU Works has an unlimited storage capacity to store any and all research data produced by the DEEPEND consortium.

Data Backup: All NSU Works production servers are maintained at a high availability colocation facility with multiple backbone connections and back-up generators. The facility is secure and requires physical tokens (badges) for access. Bepress maintains failover web, database, and storage servers to continue to serve content in case of failure. NSU Works has a real-time redundancy that runs continuously, and Bepress takes full nightly back-ups of the entire database. The back-ups are stored away from the colocation facility in a separate physical location. All of the uploaded files are stored in triplicate in a redundant storage cluster, as well as being backed up offsite to a 3rd-party cloud service that specializes in data archiving and back-up.

User Level Access: Access control can be added at the collection level or the record level and be based on IP address, email domain, or specific email address. The primary file of access controlled records will only be available to a designated group. The metadata and any supplemental content added to a record will be available open access at all times and crawled by search engines.

Support: Michele Gibney (NSU Works Repository Manager) and Keri Baker (Oceanography Collection Repository Manager), in addition to Dr. Matthew Johnston, are the main points of

contact. Designated support staff members at Bepress are available to help with any additional concerns.

NSU Works support staff:

Michele Gibney
Nova Southeastern University
8000 N Ocean Drive
Dania Beach, FL, 33004
(954)262-3692
mgibney@nova.edu

Keri Baker
Nova Southeastern University
8000 N Ocean Drive
Dania Beach, FL, 33004
(954)262- 3733
kbaker1@nova.edu

NSU Works will house all raw and in-process data (for some PIs) in order to facilitate ease of access to all consortium members. These in-process and raw data will not be publically accessible.   Access to the DEEPEND NSU Works data server will be restricted to DEEPEND personnel; however, access may also be granted to other investigators with permission from the DEEPEND Principal Investigator.

Except for specimen and water samples, all DEEPEND datasets will be transferred to the GRIIDC data archive or an appropriate U.S. national data repository, such as GenBank for genetics work, U.S. IOOS National Glider Data Assembly Center for glider data, and the Navy DoD Supercomputing Resource Center (Navy DSRC) for HYCOM model data, prior to the project end date. Upon transfer to the long-term repository, access to the datasets is expected to be unrestricted and controlled by the repository managers. After publication in GRIIDC or a U.S. national data repository, publically accessible links will be created in NSU Works pointing to the GRIIDC/repository datasets.

Other In-process Data

PI's Frank, Sutton, Lopez, Shivji, Boswell

In-process data compiled by researchers Frank, Sutton, Lopez, Shivji, and Boswell will be housed either on the NSU Works platform or on the Oceans12 data server. Both systems maintain nightly, triplicate backup to prevent data loss.

PI Hu:

In-process data will be housed on a lab PC and the lab's shared disk space on the USF Tampa servers. The servers are backed up nightly.

PI Bracken-Grissom:

In-process data will be housed on lab hard drives and backed up on FIU's high-performance 'Panther Cluster' storage system.   The Panther Cluster is backed up nightly.

PI Judkins:

In-process data will reside on the lab's shared disk space on the USF Tampa servers. The servers are backed up nightly.

PIs Penta and deRada:

In-process data will be stored at the NAVY DSRC mass storage facilities in two copies: permanent backup on tape, semi-permanent on disks/tape. Backups are performed nightly.

PI Eytan:

In-process data will reside on Amazon cloud drive.

PI Romero

In-process data will reside on a lab workstation that is backed up daily to a networked lab server.   The servers are backed up nightly.

PI's Wells & Rooker:

In-process data will reside on Google drive.

Personnel Collecting/Generating data

Researcher	Role
Bernard, Andrea	Postdoctoral Researcher
Boswell, Kevin	Co-PI
Bracken-Grissom, Heather	Co-PI
Bush, Stephanie	Research Scientist
Cook, April	Research Scientist
deRada, Sergio	Research Scientist
Easson, Cole	Postdoctoral Researcher
English, David	Research Scientist
Eytan, Ron	Co-PI
Fenolio, Danté	Co-PI
Frank, Tamara	Co-PI
Hartland, Richard	Student
Hu, Chuanmin	Co-PI
Johnston, Matt	Research Scientist
Judkins, Heather	Co-PI
Kovach, Charles	Research Scientist

Lembke, Chad	Research Scientist
Lopez, Joe	Co-PI
Moore, Jon	Co-PI
Nizinski, Martha	Co-PI
Penta, Brad	Co-PI
Pérez-Moreno, Jorge Luis	Student
Pruzinski, Nina	Student
Romero, Isabel	Co-PI
Rooker, Jay	Co-PI
Shivji, Mahmood	Co-PI
Sutton, Tracey	PI
Timm, Laura	Student
Vazquez Miranda, Hernán	Postdoctoral Researcher
Warren, Joe	Research Scientist
Weber, Max	Student
Weber, Megan	
Wells, David	Co-PI
White, James	Research Scientist
Youngbluth, Marsh	Co-PI

Section 2: Detailed Task Information

The DEEPEND consortium will conduct research organized in five main task areas. Below is a description of each task and associated datasets. The estimated size is indicated **per cruise** and we anticipate a total of **six cruises.**

Task Title: Taxonomy/Community Ecology

Task Description: A panel of deep-sea taxonomic experts will team with five genetics laboratories to produce the most comprehensive inventory to date of the biodiversity and community structure of the Gulf oceanic fauna, with emphasis on pelagic fishes, crustaceans, molluscs, gelatinous zooplankton, and microbes. This inventory will be based on literature records, past sampling (2010-11) and future sampling (2015-17).

Data descriptions for this task, organized by dataset:

PI Name: Dr. Tracey Sutton **Contact:** tsutton1@nova.edu **Institution:** Nova Southeastern University	**UDIs:** R4.x257.226:0001,R4.x257.226:0004 R4.x257.226:0005,R4.x257.226:0002 R4.x257.226:0006,R4.x257.226:0007 R4.x257.226:0008,R4.x257.226:0009 R4.x257.226:0010,R4.x257.226:0003

	R4.x257.226:0011,R4.x257.226:0012 R4.x257.226:0013,R4.x257.226:0014, R4.x257.226:0015
What data will be generated in the research	• Ship data • MOCNESS (10 m^2 Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS or MOC10)) data • Biological data (counts, lengths, weights, etc.) • CTD data • Photos/Videos - not including Dante's work • **Estimated size (per cruise): 10GB**
What data types will you be creating or capturing	.ascii, .raw, .accdb, .xls, .doc, .txt, .wmv, .pdf, .jpeg, .tiff, etc.
How will you capture or create the data	• Logbooks transcribed into excel/access • Direct entry into excel/access • MOCNESS SENSORS CTD sensors • Cameras
When will these data be generated?	Per cruise basis and then updated or processed post-cruise
How long will each dataset be kept locally and why?	• The shipboard data (logbooks), MOCNESS data, Photos/Videos, and CTD data can be uploaded to GRIIDC shortly after the cruise (some processing required). • The biological database will be evolving as organisms are identified to the lowest taxonomic level. This database will not be kept locally (on the Oceans12 data server or NSU Works) longer than 1 year before submission to GRIIDC.
Will any existing data be used in this research task?	We may use some data collected by the NRDA cruises to form our species database or to compare those data to the current dataset. These data will not be shared through GRIIDC.

Co-PI Name: Dr. Tamara Frank Contact: tfrank1@nova.edu Institution: Nova Southeastern University	UDIs: R4.x257.226:0001,R4.x257.226:0004 R4.x257.226:0005
What data will be generated in the research	• Quantification of diversity and abundance of crustacean micronekton species • Correlation between crustacean community composition and environmental parameters (temperature, salinity, wet season vs. dry season, distance from land, bottom depth, O$_2$ concentration)

	• Day and night distribution patterns • Life history – reproduction and ontogenetic migrations • Voucher specimens • Image database of voucher specimens and their key characteristics • **Estimated size (per cruise): 100GB**
What data types will you be creating or capturing	• Excel and access spreadsheets and databases • Images • Voucher specimens
How will you capture or create the data	• Animals will be caught on research cruises with a MOC-10 or a 3 m^2 Tucker Trawl • Samples will be roughly sorted on shipboard and stored in 4% formalin, RNAlater, or 95% ethanol (vouchers) • Identification to species will usually take place in land-based laboratories • Data will be stored on datasheets, transcribed daily into excel databases and backed up on the NSU Works data repository • Images will be taken with the aid of a microscope camera
When will these data be generated?	• Starting from the first cruise in May 2015, through Year 3
How long will each data be kept locally and why?	• All non-voucher specimens that have been quantified to species, weighed and measured will be kept in the PI's lab until all manuscripts resulting from these data have been published, at which time the common species will be offered to members of the consortium first, and then on other teaching collections. Three years after publication, all remaining common species will be disposed of. Rare species will be kept indefinitely in secure storage • Voucher specimens will be offered to museum repositories, and those that are not sent to museums will be kept indefinitely in secure storage • Datasheets will be kept indefinitely in the PIs lab, housed on the Oceans12 data server. • At the time of publication, spreadsheet and image databases will be made available to the public through GRIIDC.
Will any existing data be used in this research task?	• Once they become publically available, data from the NRDA cruises that occurred right after the DWH oil spill will be used as a reference/baseline

	against which changes or lack thereof in the crustacean micronektonic community composition can be monitored. These data will not be shared through GRIIDC.

Co-PI Name: Dr. Jose Lopez Contact: joslo@nova.edu Institution: Nova Southeastern University	UDIs: R4.x257.228:0001, R4.x257.228:0008, R4.x257.228:0009
What data will be generated in the research?	• Microbiome DNA or RNA sequence data • Long text files • Tab-delimited text (Excel) • Proprietary Fastq/fasta sequence files • **Estimated size (per cruise): 50-1000GB**
What data types will you be creating or capturing?	Experimental measures
How will you capture or create the data?	High-throughput sequencing • Illumina • MiSeq DNA sequencer • HiSeq DNA sequencer
When will these data be generated?	1-2 months after each cruise for every year of the project
How long will each data set be kept locally and why?	Raw and in-process data will be stored and backed up on the NSU Works data repository or Oceans12 data server for the duration of the project and at least 5 years afterward. After the data are processed and QC'd, they will be deposited in public databases (e.g. GenBank) and linked to in GRIIDC and published
Will any existing data be used in this research task?	No historical data will be used

Task Title: Organismal Ecology

Task Description: A variety of traditional and biochemical approaches will be used to: 1) investigate past and presence hydrocarbon contamination in the oceanic fauna; 2) characterize the food web structure of the oceanic Gulf as it relates to pathways of contamination transfer across trophic levels; and 3) investigate the long-term consequences of the spill on the recruitment of economically valuable oceanic taxa (e.g., tunas, billfishes).

Data descriptions for this task:

Co-PI Name: Dr. Isabel Romero Contact: isabelromero@mail.usf.edu	UDIs: R4.x257.227:0001, R4.x257.227:0016, R4.x257.227:0017

Institution: University of South Florida	
What data will be collected/generated/produced/acquired through this research task?	• Composition and quantification of polycyclic aromatic hydrocarbons in deep-pelagic fauna • **Estimated size (per cruise): 5GB**
What data types will you be creating or capturing?	• Field deep-pelagic fauna samples • Experimental measures
How will you capture or create the data?	Chemical analysis will be generated by GC-MS instrument
When will these data be generated?	Data will be generated throughout Year 1, 2, and 3.
How long will each data set be kept locally and why?	Raw data will be stored and backed up on the NSU Works data repository and also will be kept locally for the duration of the project, for quality control purposes
Will any existing data be used in this research task?	Historical contaminant data will be used to establish the baseline. These data will not be shared through GRIIDC.

Co-PI Name: Dr. Jay Rooker **Contact:** rookerj@tamug.edu **Institution:** Texas A&M University at Galveston	**UDIs:** R4.x257.227:0002, R4.x257.227:0012, R4.x257.227:0010, R4.x257.227:0011, R4.x257.227:0013
What data will be collected/generated/produced/acquired through this research task?	• Larval fishes from 48 stations in N GoM. • Cruise data collected at each site including Lat/Lon, Date, Time, and environmental (Temp, Salinity, DO). • Photographs of selected specimens. **Estimated size (per cruise): 1GB**
What data types will you be creating or capturing?	Presence/absence, density data, and assemblage/community measurements for all pelagic fish taxa collected in excel spreadsheets
How will you capture or create the data?	Presence/absence and density data captured through sorting of collection from neuston and bongo net gears; visual and genetic identification (PCR assays)
When will these data be generated?	Data generated throughout each of the three years; collections/processing of each summer survey completed within the year sampled.
How long will each data set be kept locally and why?	Data will be stored at TAMUG for the duration of the project and backed up regularly on NSU Works Data Repository

	Historical TAMUG ichthyoplankton cruise data 2006-2009 will be used as our baseline. Also, post-spill ichthyoplankton surveys (2010-2012) will be used with the current dataset. These data will not be shared through GRIIDC.
Will any existing data be used in this research task?	

Co-PI Name: Dr. Jay Rooker **Contact:** rookerj@tamug.edu **Institution:** Texas A&M University at Galveston	**UDIs:** R4.x257.227:0003,R4.x257.227:0014 R4.x257.227:0015
What data will be collected/generated/produced/acquired through this research task?	• Chemistry data on otoliths (ear stones) of deepwater fishes in GoM. • Trace element (Li, Mg, Ba, Sr, Mn) and stable isotope data (C and O) • **Estimated size (per cruise): 10 MB**
What data types will you be creating or capturing?	Otolith chemistry data in excel spreadsheets
How will you capture or create the data?	Trace element and stable isotope data captured on inductively coupled plasma mass spectrometer and stable isotope mass spectrometer, respectively
When will these data be generated?	Otolith processed each of the three years following the cruises (fall/winter) and available by spring
How long will each data set be kept locally and why?	Data will be stored at TAMUG for the duration of the project and backed up regularly on NSU Works Data Repository
Will any existing data be used in this research task?	No historical data will be used

Co-PI Name: Dr. David Wells **Contact:** wellsr@tamug.edu **Institution:** Texas A&M University at Galveston	**UDIs:** R4.x257.227:0004,R4.x257.227:0006, R4.x257.227:0008,R4.x257.227:0005, R4.x257.227:0009,R4.x257.227:0008
What data will be collected/generated/produced/acquired through this research task?	• Feeding ecology of model species in food webs of the Gulf of Mexico. Specifically, data such as stomach contents and stable isotopes will be collected. • **Estimated size (per cruise): 1 GB**
What data types will you be creating or capturing?	Stomach content and stable isotope data in excel spreadsheets.

How will you capture or create the data?	Stomach content data will be captured through sorting of the stomachs while stable isotope data will be acquired through a stable isotope mass spectrometer.
When will these data be generated?	Data will be generated throughout each of the three years of collection and processing.
How long will each data set be kept locally and why?	Raw data will be stored and backed up on the NSU Works data repository and also will be kept locally for the duration of the project, for quality control purposes.
Will any existing data be used in this research task?	Historical stable isotope data will be used to establish the baseline. These data will not be shared through GRIIDC.

Task Title: Genetic Diversity and Connectivity

Task Description: In addition to molecular taxonomy (i.e., DNA "barcoding"), genetic and genomic techniques will be used to assess population connectivity within different regions of the Gulf, as well as investigating the potential effect of the *Deepwater Horizon* oil spill on genetic diversity/population structure.

Data descriptions for this task:

Co-PI Name: Dr. Heather Bracken-Grissom **Contact:** heather.bracken@gmail.com **Co-PI Name:** Dr. Heather Judkins **Contact:** judkins@mail.usf.edu **Institution:** Florida International University, USF St. Petersburg	**UDIs:** R4.x257.228:0003,R4.x257.228:0006, R4.x257.228:0007,R4.x257.228:0004
What data will be collected/generated/produced/acquired through this research task?	• The raw data files will be fastq (genetic diversity) and ab1 files (barcoding) or fasta files. • **Estimated size (per cruise): 10GB**
What data types will you be creating or capturing?	• All DNA/RNA sequence data including individual sequences and alignments • GenBank numbers will be provided.
How will you capture or create the data?	• DNA/RNA sequencing techniques. • Platform to be used is ABI 3730xl DNA Analyzer 96-capillary automated sequencer or HiSeq 2500 Illumina Next Generation Sequencing platform

When will these data be generated?	After samples are collected from the research cruises and brought back to individual labs, years 1, 2, 3
How long will each data set be kept locally and why?	All in-process datasets will be kept locally until results are published, analyses are run, or it has been 1 year. Raw data will be stored and backed up on the NSU Works data repository.
Will any existing data be used in this research task?	All data will be newly generated

Co-PI Name: Dr. Ron I. Eytan **Contact:** ron.eytan@gmail.com **Institution:** Texas A&M University at Galveston	**UDIs:** R4.x257.228:0002,R4.x257.228:0010, R4.x257.228:0011,R4.x257.228:0005
What data will be collected/generated/produced/acquired through this research task?	• The raw data files will be fastq (genetic diversity) and ab1 files (barcoding) or fasta files. • **Estimated size (per cruise): 2GB**
What data types will you be creating or capturing?	• All DNA/RNA sequence data including individual sequences and alignments • GenBank numbers will be provided.
How will you capture or create the data?	• DNA/RNA sequencing techniques • Platform to be used is ABI 3730xl DNA Analyzer
When will these data be generated?	After samples are collected from the research cruises and brought back to individual labs, years 1, 2, 3
How long will each data set be kept locally and why?	All in-process datasets will be kept locally until results are published, analyses are run, or it has been 1 year. Raw data will be stored and backed up on the NSU Works data repository. After the data is processed and QC'd, it will be deposited in public databases (e.g. GenBank) and linked to in GRIIDC and published.
Will any existing data be used in this research task?	All data will be newly generated

Co-PI Name: Dr. Mahmood Shivji **Contact:** mahmood@nova.edu **Institution:** Nova Southeastern University	**UDIs:** R4.x257.228:0002,R4.x257.228:0010, R4.x257.228:0011,R4.x257.228:0005
What data will be collected/generated/produced/acquired through this research task?	• The raw data files will be fastq (genetic diversity) and ab1 files (barcoding) or fasta files. • **Estimated size (per cruise): 2GB**

What data types will you be creating or capturing?	• All DNA/RNA sequence data including individual sequences and alignments • GenBank numbers will be provided.

How will you capture or create the data?	• DNA/RNA sequencing techniques
	Platform to be used:
	   o ABI 3730xl DNA Analyzer 96-capillary automated sequencer,
	   o ABI 3130 DNA Analyzer 4-capillary automated sequencer
	   o HiSeq 2500 Illumina Next Generation Sequencing platform

When will these data be generated?	After samples are collected from the research cruises and brought back to individual labs, years 1, 2, 3
How long will each data set be kept locally and why?	All datasets will be kept locally on the NSU Works or Oceans12 data server until results are published, analyses are run, or it has been 1 year. Raw data will be stored and backed up on the NSU Works data repository. After the data is processed and QC'd, it will be deposited in public databases (e.g. GenBank) and linked to in GRIIDC and published.
Will any existing data be used in this research task?	All data will be newly generated

Task Title: Acoustics

Task Description: High-resolution acoustic sensing will be directed sampling to characterize the distribution and abundance of the oceanic fauna across large spatial scales (km).

Data descriptions for this task:

PI Name: Dr. Kevin Boswell **Contact:** kevin.boswell@fiu.edu **Institution:** Florida International University	**UDIs:** R4.x257.229:0001,R4.x257.229:0002, R4.x257.229:0003
What data will be collected/generated/produced/acquired through this research task?	• Raw acoustic scattering data of organisms from the water column • Processed data flows of acoustic data from each frequency • Processed water-column level CTD data • Summary data from final acoustic processing • **Estimated size (per cruise): 10GB**
What data types will you be creating or capturing?	• Processed acoustic scattering profiles • Echogram images • metrics describing the acoustically detected

	patterns
How will you capture or create the data?	Data will be generated from processing within software- Echoview and Matlab
When will these data be generated?	Following each cruise
How long will each data set be kept locally and why?	In-process data will be kept locally in perpetuity and backed up weekly in triplicate. Raw data will be uploaded to the NSU Works repository.
Will any existing data be used in this research task?	We will reference acoustic data collected aboard ships used during NRDA process, specifically 2010-2011. These data will not be shared through GRIIDC.

Task Title: Physical Oceanography

Task Description: Advanced in situ (e.g., underwater gliders) and remote (satellite) sensing will be combined with biophysical modelling to determine the drivers responsible for oceanic community distribution and dynamics.

Data descriptions for this task:

Co-PI Name: Dr. Chuanmin Hu **Contact:** huc@usf.edu **Institution:** University of South Florida, College of Marine Science	**UDIs:** R4.x257.230:0001,R4.x257.230:0015, R4.x257.230:0016,R4.x257.230:0002, R4.x257.230:0017,R4.x257.230:0018, R4.x257.230:0003,R4.x257.230:0019, R4.x257.230:0020,R4.x257.230:0004, R4.x257.230:0010,R4.x257.230:0011, R4.x257.230:0012,R4.x257.230:0013, R4.x257.230:0014,R4.x257.230:0005, R4.x257.230:0021,R4.x257.230:0022, R4.x257.230:0006,R4.x257.230:0023, R4.x257.230:0024
What data will be collected/generated/produced/acquired through this research task?	**From Gliders:** • Glider time-series contains location, time, & depth with values from several sensors • These often include conductivity, temperature, depth, chl_Fluor, CDOM_Fluor, and sometimes 4 chan. irradiance data **Collected at Sea:** • Optical absorption spectra for particulate and dissolved matter • Possible near-surface, vertical profiles of

	multispectral optical backscattering, Chl_a, and CDOM fluorescence • Shipboard remote sensing reflectance spectra (conditions permitting) **From Satellite:** • MODIS chl. and temperature imagery • **Estimated size (per cruise): 30GB**
What data types will you be creating or capturing?	• Glider data is received as binary files, but these are converted to arrays of sensor measurements (NetCDF and other formats). • Data collected from the ship optical absorption and reflection spectra and arrays of sensor measurements. • Data from satellite measurements will be multi-banded imagery (HDF, NetCDF, or various formats are used for exchange of imagery.
How will you capture or create the data?	**At Sea:** • Discrete samples collected at sea using filtration, later extraction & processing at USF • Instrument data downloaded at sea, initial values generated, final values produced after incorporating multiple sensors streams and ancillary info. **From Satellite:** • MODIS raw data downloaded from NASA, and processed at USF to appropriate products
When will these data be generated?	A subset of the glider data becomes available during each deployment, with the complete glider dataset becoming available within a few weeks of glider recovery. Shipboard data are usually processed within weeks to months, satellite data are usually available the same day
How long will each data set be kept locally and why?	Glider data are stored locally, and then transmitted to U.S. IOOS National Glider Data Assembly Center. Raw shipboard data are stored at the PI's and processed data submitted to an appropriate database. Raw satellite data are stored at NASA, derived imagery is kept at the PI's lab and/or made available via USF website.

Will any existing data be used in this research task?	All data will be newly generated

Co-PI Name: Dr. Bradley Penta **Contact:** brad.penta@nrlssc.navy.mil **Co-PI Name:** Sergio deRada **Contact:** sergio.derada@nrlssc.navy.mil **Institution:** Naval Research Laboratory	**UDIs:** R4.x257.230:0007,R4.x257.230:0008, R4.x257.230:0009
What data will be collected/generated/produced/acquired through this research task?	• Numerical simulation model output • **Estimated size (per cruise): 1.5TB – total 9TB regardless of total number of cruises**
What data types will you be creating or capturing?	COARDS compatible netCDF
How will you capture or create the data?	Computer simulation
When will these data be generated?	Before and during cruises
How long will each data set be kept locally and why?	Data will be kept locally as space permits and at the Navy DoD Supercomputing Resource Center (Navy DSRC) indefinitely
Will any existing data be used in this research task?	Historical remote sensing and model data will be used to initialize and force the model. These data will not be shared through GRIIDC.

7. Documentation

Sample naming convention:

All samples collected during the cruises will be named according to the following conventions:

Cruise Name/Number – DDMMMYY – Gear – Station Solar Cycle - Trawl# - Net#

Example: DP01-01May15-MOC10-B250N-003-N4

Cruise Name/Number: The abbreviation for DEEPEND cruises will be DP followed by a 2 digit cruise number (starting with '01')
Date: The date format will be a 2 digit day, three letter month abbreviation, 2 digit year (ex. '02May15')

Gear: The following gear abbreviations will be used-

Abbrev.	Gear
MOC10	10 meter MOCNESS Net
TT	Tucker Trawl
NN	Neuston Net
BN	Bongo Net
CTD	CTD Water Samples

Station: We will use the existing SEAMAP stations
Solar Cycle: Day or Night (D or N)
Trawl#: We will sequentially number each trawl for the duration of the DEEPEND project using a 3 digit number (starting with 001)
Net#: This only applies to the MOC10 samples and refers to the net from which the sample was collected. We will use the letter N and a single digit number beginning with N0 following this scheme:

Net#	Depth Range
N0	Surface-1500 m
N1	1500 m-1200 m
N2	1200 m-1000 m
N3	1000 m-600 m
N4	600 m-200 m
N5	200 m-Surface

DEEPEND Cruise Tissue Identification

Trawl# Taxa Net#– Tissue Tag#

Example: 020C2-00254

Trawl#: We will sequentially number each trawl for the duration of the DEEPEND project using a 3 digit number (starting with 001)
Taxa: This will ensure a unique identification for each genetic sample between different taxa within the same trawl. We will follow:

Abbrev.	Taxa
C	Crustacea
F	Fish
M	Mollusca

Net#: This only applies to the MOC10 samples and refers to the net from which the sample was collected. See section above for key.

Tissue Tag#: This will consist of a 5 digit number assigned by the geneticist taking the tissue samples and will be placed with both the tissue sample and the voucher specimen.

DEEPEND Cruise Water Filter Samples

> **Cruise Name/Number – DDMMMYY – Gear – Station Solar Cycle – Cast# - Depth - Lab**

Example: DP01-02May15-CTD-B250D-012-1200m-USF

Cast#: Each CTD cast will be sequentially numbered for the duration of the DEEPEND project using a 3 digit number (starting with 001)
Depth: The depth from which the water was collected and the unit of measure.
Lab: The laboratory in which the sample will be analyzed.

DEEPEND Cruise Microbe Genetic Samples

> **Gear Cast# - Depth – Lab**

Example: CTD012-1200m-NSU

See explanation of sample code above.

File naming convention:

All files for a dataset will be prefaced by the **cruise name/number** – i.e. DP01 for cruise #01, etc.

8. Ethics and Compliance

There are no ethical issues or privacy issues, copyright restrictions, or IRB/HIPPA issues anticipated for any DEEPEND dataset. DEEPEND datasets will not need to be anonymized in any way. If such issue should present, the PI, Dr. Tracey Sutton, will manage the concerns to ensure full compliance with Federal and State regulations.

9. Storage and Backup

Please see section 1.6 for our data storage and backup plan.

10. Archives

Genetic datasets (Co-PIs Shivji, Eytan, Bracken-Grissom, and Lopez) will be deposited in NCBI GenBank and linked to in GRIIDC.   Glider data (Co-PI Hu) will be deposited in the U.S. IOOS National Glider Data Assembly Center databased and linked to in GRIIDC.   HYCOM model data (Co-PI deRada and Penta) will be deposited in the Navy DoD Supercomputing Resource Center (Navy DSRC) and linked to in GRIIDC.

All other data will be deposited in the GRIIDC database.

Section 3: Dataset Information Forms

All finalized data will be housed in GRIIDC in the following datasets and we will create public links to these datasets from within NSU Works. Existing data from other projects will not be included in or shared through GRIIDC.

Task Title	Dataset Title (in GRIIDC)	UDI
Taxonomy / Community Ecology	DEEPEND: Inventory of Gulf oceanic fauna data including species, weight, and measurements. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.226:0001
	DEEPEND: Inventory of Gulf oceanic fauna data including species, weight, and measurements. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.226:0004
	DEEPEND: Inventory of Gulf oceanic fauna data including species, weight, and measurements. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.226:0005
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP01, May 2015	R4.x257.226:0002
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP02, August 2015	R4.x257.226:0006
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP03, May 2016	R4.x257.226:0007
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP04, August 2016	R4.x257.226:0008
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP05, May 2017	R4.x257.226:0009
	DEEPEND: Multiple Opening Closing Net Environmental Sensing System (MOCNESS) instrument data. Cruise DP06, August 2017	R4.x257.226:0010
	DEEPEND: Ship data - M/V Point Sur. Cruise DP01, May 2015	R4.x257.226:0003
	DEEPEND: Ship data - M/V Point Sur. Cruise DP02, August 2015	R4.x257.226:0011
	DEEPEND: Ship data - M/V Point Sur. Cruise DP03, May 2016	R4.x257.226:0012
	DEEPEND: Ship data - M/V Point Sur. Cruise DP04, August 2016	R4.x257.226:0013
	DEEPEND: Ship data - M/V Point Sur. Cruise DP05, May 2017	R4.x257.226:0014
	DEEPEND: Ship data - M/V Point Sur. Cruise DP06, August 2017	R4.x257.226:0015

	DEEPEND: Composition and quantification of Polycyclic Aromatic Hydrocarbons in deep-pelagic fauna. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.227:0001
	DEEPEND: Composition and quantification of Polycyclic Aromatic Hydrocarbons in deep-pelagic fauna. Cruises DP03 and DP04 from January 2016 - December 2017	R4.x257.227:0016
	DEEPEND: Composition and quantification of Polycyclic Aromatic Hydrocarbons in deep-pelagic fauna. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.227:0017
	DEEPEND: Presence/absence and density data for epipelagic trolls from 48 stations in the northern Gulf of Mexico, including cruise data. Cruises LF2015A and LF2015B from June 2015 - July 2015	R4.x257.227:0002
	DEEPEND: Presence/absence and density data for epipelagic trolls from 48 stations in the northern gulf of Mexico, including cruise data. Cruises LF2016A and LF2016B from June 2016 - July 2016	R4.x257.227:0012
	DEEPEND: Presence/absence and density data for epipelagic trolls from 48 stations in the northern gulf of Mexico, including cruise data. Cruises LF2017A and LF2017B from June 2017 - July 2017	R4.x257.227:0010
	DEEPEND: Historical larval catch dataset before (2006–2009), during (2010), and after the Deep Water Horizons Oil Spill (2011-2013)	R4.x257.227:0011
Organismal Ecology	DEEPEND: Life history and ecosystem parameter data from before (2007-2009) and well after (2015-2017) the Deep Water Horizons Oil Spill	R4.x257.227:0013
	DEEPEND: Chemistry data on otoliths of deepwater fishes in the Gulf of Mexico. Cruises LF2015A and LF2015B from June 2015 - July 2015	R4.x257.227:0003
	DEEPEND: Chemistry data on otoliths of deepwater fishes in the Gulf of Mexico. Cruises LF2016A and LF2016B from June 2016 - July 2016	R4.x257.227:0014
	DEEPEND: Chemistry data on otoliths of deepwater fishes in the Gulf of Mexico. Cruises LF2017A and LF2017B from June 2017 - July 2017	R4.x257.227:0015
	DEEPEND: Stomach content analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2015A and LF2015B from June 2015 - July 2015	R4.x257.227:0004
	DEEPEND: Stomach content analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2016A and LF2016B from June 2016 - July 2016	R4.x257.227:0006
	DEEPEND: Stomach content analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2017A and LF2017B from June 2017 - July 2017	R4.x257.227:0008

	DEEPEND: Stable isotope analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2015A and LF2015B from June 2015 - July 2015	R4.x257.227:0005
	DEEPEND: Stable isotope analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2016A and LF2016B from June 2016 - July 2016	R4.x257.227:0009
	DEEPEND: Stable isotope analyses of model species to examine marine food webs of the Gulf of Mexico. Cruises LF2017A and LF2017B from June 2017 - July 2017	R4.x257.227:0008
Genetic Diversity and Connectivity	DEEPEND: Microbiome and bacterioplankton rRNA gene sequence data collected from Gulf of Mexico seawater samples. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.228:0001
	DEEPEND: Microbiome and bacterioplankton rRNA gene sequence data collected from Gulf of Mexico seawater samples. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.228:0008
	DEEPEND: Microbiome and bacterioplankton rRNA gene sequence data collected from Gulf of Mexico seawater samples. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.228:0009
	DEEPEND: DNA sequences of the mitochondrial Cytochrome c oxidase I (COI) genes from deep sea fishes. Cruises DP01 and DP02 from May 2015 - August 2015	R4.x257.228:0002
	DEEPEND: DNA sequences of the mitochondrial Cytochrome c oxidase I (COI) genes from deep sea fishes. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.228:0010
	DEEPEND: DNA sequences of the mitochondrial Cytochrome c oxidase I (COI) genes from deep sea fishes. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.228:0011
	DEEPEND: DNA sequences and genome-scale single nucleotide polymorphisms of deep sea fishes from the Gulf of Mexico. Cruises DP01 to DP06 from May 2015 - December 2017	R4.x257.228:0005
	DEEPEND: 16S and COI barcoding sequences for crustaceans and cephalopods collected from the Northern Gulf of Mexico for Cruises DP01 and DP02 from May 2015 - August 2015	R4.x257.228:0003
	DEEPEND: 16S and COI barcoding sequences for crustaceans and cephalopods collected from the Northern Gulf of Mexico for Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.228:0006
	DEEPEND: 16S and COI barcoding sequences for crustaceans and cephalopods collected from the Northern Gulf of Mexico for Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.228:0007

	DEEPEND: DNA/RNA sequence data used to measure crustacean and cephalopod genetic diversity collected from the Northern Gulf of Mexico for cruises DP01 to DP05 from May 2015 - December 2017	R4.x257.228:0004
Acoustics	DEEPEND: Raw acoustic scattering data of organisms from the water column. Cruises DP01 and DP02 from May 2015 - August 2015	R4.x257.229:0001
	DEEPEND: Raw acoustic scattering data of organisms from the water column. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.229:0002
	DEEPEND: Raw acoustic scattering data of organisms from the water column. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.229:0003
Physical Oceanography	DEEPEND: Chlorophyll concentration and optical absorption spectra of particulate and dissolved material, collected at several depths for 6 Coastal Florida stations. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.230:0001
	DEEPEND: Chlorophyll concentration and optical absorption spectra of particulate and dissolved material, collected at several depths for 6 Coastal Florida stations. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0015
	DEEPEND: Chlorophyll concentration and optical absorption spectra of particulate and dissolved material, collected at several depths for 6 Coastal Florida stations. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0016
	DEEPEND: Spectral remote sensing reflectance of the water surface collected at multiple stations. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.230:0002
	DEEPEND: Spectral remote sensing reflectance of the water surface collected at multiple stations. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0017
	DEEPEND: Spectral remote sensing reflectance of the water surface collected at multiple stations. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0018
	DEEPEND: Near surface fluorescence and optical scattering profiles at CTD sampling stations. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.230:0003
	DEEPEND: Near surface fluorescence and optical scattering profiles at CTD sampling stations. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0019
	DEEPEND: Near surface fluorescence and optical scattering profiles at CTD sampling stations. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0020
	DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP01, May 2015	R4.x257.230:0004

DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP02, August 2015	R4.x257.230:0010
DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP03, May 2016	R4.x257.230:0011
DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP04, August 2016	R4.x257.230:0012
DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP05, May 2017	R4.x257.230:0013
DEEPEND: Conductivity, temperature and depth (CTD) data for DEEPEND stations. Cruise DP06, August 2017	R4.x257.230:0014
DEEPEND: Salinity, temperature, and fluorescence measurements from an oceanographic glider deployed in the North Eastern Gulf of Mexico. Cruise DP02 - August 2015	R4.x257.230:0005
DEEPEND: Salinity, temperature, and fluorescence measurements from an oceanographic glider deployed in the North Eastern Gulf of Mexico. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0021
DEEPEND: Salinity, temperature, and fluorescence measurements from an oceanographic glider deployed in the North Eastern Gulf of Mexico. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0022
DEEPEND: Satellite derived ocean color and temperature imagery for the North Eastern Gulf of Mexico. Cruises DP01 and DP02 from May 2015-August 2015	R4.x257.230:0006
DEEPEND: Satellite derived ocean color and temperature imagery for the North Eastern Gulf of Mexico. Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0023
DEEPEND: Satellite derived ocean color and temperature imagery for the North Eastern Gulf of Mexico. Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0024
DEEPEND: Numerical simulation of the Gulf of Mexico (Gom) using the Hybrid Coordinate Ocean Model (HYCOM). Cruises DP01 and DP02 from May 2015 - December 2015	R4.x257.230:0007
DEEPEND: Numerical simulation of the Gulf of Mexico (Gom) using the Hybrid Coordinate Ocean Model (HYCOM). Cruises DP03 and DP04 from January 2016 - December 2016	R4.x257.230:0008
DEEPEND: Numerical simulation of the Gulf of Mexico (Gom) using the Hybrid Coordinate Ocean Model (HYCOM). Cruises DP05 and DP06 from January 2017 - December 2017	R4.x257.230:0009