**Appendix 1:**

**Data management plan: Master DMP for the NMFS Optics Strategic Initiative**

**Reference to Master DM Plan (if applicable)**

*Data Management Plans (DMP) may be hierarchical. If this DMP inherits provisions from a higher-level DMP already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DMP.*

URL of higher-level DMP (if any) as submitted to DM Plan Repository:

**Introduction and Overview:**

* The Optics Strategic Initiative is funding projects to accelerate the transition of end-to-end automated processing pipelines to operational status across a wide variety of NMFS programs. To do this we will invest in upgrades to optical hardware, processing software, and a set of projects that will capitalize on those investments by pushing them to operational status TRL 9 by 2026. To do so requires data storage capacity located on-premises servers or via cloud infrastructure for existing labeled image libraries and collected imagery (or both). Automation also requires access to high-speed computing (i.e. access to GPU) and appropriate software through which pipelines flow.
* To facilitate this each data collection project was requested to include cloud computing costs into their budgets given the knowledge there will not be support from sources external to the SI’s. Individual projects will also be required to provide local on-premises resources to to handle data following acquisition and transfer to local networking, cloud storage, and backup copies created and archived.
* We preferentially selected projects with high TRL readiness levels, and in doing so, assured that selected transition projects already have data resources in place that their individual budgets will then supplement and push infrastructure and processing to cloud resources.
* We selected Core projects that must be funded in full in order for the SI to be successful including upgrades to existing software. Thus the pipelines are fully funded in all funding scenarios requested by NMFS Science Board. Thus the data is largely dependent on the number of projects we can fund under the three funding scenarios.
* Projects were selected from 1 page project narratives with high level budget information. Projects were distilled and efficiencies sought during review and selection. However these projects were not requested to go through the additional step of creating a data management plan. Individual plans will be created with transition plans for the selected projects. Some projects with existing transition plans will also have existing data management plans in place.
* We intend for this to serve as the Master DMP for the OSI project.

1. **General Description of Data to be Managed**
   1. Summary description of the data:

Optical data will be, or have been collected, across a broad range of projects but in general could be classified for the cases of the SI into Fish, Plankton, Benthos, Aerial. These use cases were defined because the underlying computer science problem (e.g. segmentation of corals) had similar properties. Thus processing pipelines can be developed for those applications and thus be maximally useful across the set of transition projects funded by the SI.

Some imagery will likely be georeferenced and with proper labels made ready for use in online portals (e.g. FathomNet, CoralNET).

Photographs scale drastically in terms of the resolution of pictures necessary for a particular workflow (e.g. corals and aerial imaging require high resolution). Thus onboard demand, transfer, processing, and archiving will vary by project type and cost and demand depends on what process is being applied to the data (e.g. storage and processing in cloud environments).

Image libraries used in developing automated image algorithms also contain files for location and labeling information. These must be coupled and static as data pipelines will be dependent on understanding this structure.

Data will also include automated annotation output generated during processing steps. These data will likely be distributed across a variety of buckets determined by status (e.g. raw vs qa/qc).

* 1. Is this a one-time data collection, or an ongoing series of measurements?

Projects selected by the Optical Working group are operational projects intended to collect long-term time-series data. If successful, and desired, support will have to be sought to extend these end-to-end pipelines past the funding cycle (FY26) and will likely require an enterprise level solution from NMFS-OCIO (or other).

* 1. Actual or planned temporal coverage of the data:

Current funding will limit data management between FY23 and FY26.

* 1. Actual or planned geographic coverage of the data:

Projects span across most, if not all, Science Centers. Thus the geographic scope of the OSI projects includes all U.S. Territorial Seas in the EEZ. OSI projects could also include U.S. overseas territories in the Caribbean and South Pacific.

* 1. Type(s) of data:  
     *(e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.)*

Optical (e.g. photo, video, satellite), digital numeric, metadata.

* 1. Data collection method(s) and Standards  
     *(e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)*
     1. Detail the data collection methods, tools, and data types

As the OSI strategy was to look for solutions across a broad sweep of needs in the agency, we sought to identify gaps that address many needs across groups. Thus the OSI project data collections span a wide variety of platforms potentially including: satellite, aerial, shipbased, autonomous platforms, and animal-borne.

* + 1. Establish standardized data formats, metadata, and naming conventions to ensure data compatibility

As with the platforms themselves data standards will be project specific depending on specific needs. For instance coral segmentation models require high resolution imagery, that would be wasteful for other applications than need lower resolution.

* 1. If data are from a NOAA Observing System of Record,15 indicate name of system: NA
     1. If data are from another observing system, please specify: NA
  2. Data Storage, Security, and Privacy
     1. Describe data storage infrastructure and security measures to protect sensitive or confidential data shared between organizations

Projects are largely insular to NOAA Fisheries. Thus data storage and computing, regardless if it is on-premises or in cloud computing will be required to be in compliance with guidelines set for by the agency. Each laboratory's capacity will be Science Center and likely even Laboratory specific. Thus it’s otherwise difficult to describe as there are as many systems as projects, laboratories, and centers. We will work with Agency professionals to ensure proper handling of data.

* + 1. Address data access controls and mechanisms to safeguard sensitive information and ensure compliance with data protection and privacy regulations

Projects are largely insular to NOAA Fisheries. Thus data storage and computing, regardless if it is on-premises or in cloud computing will be required to be in compliance with guidelines set for by the agency. Each laboratory's capacity will be Science Center and likely even Laboratory specific. Thus it’s otherwise difficult to describe as there are as many systems as projects, laboratories, and centers. We will work with Agency professionals to ensure proper handling of data.

1. **Point of Contact for this Data Management Plan (author or maintainer)**
   1. Name: Matthew D. Campbell
   2. Title: Gulf and Caribbean Reef Fish Branch Chief - Southeast Fisheries Science Center
   3. Affiliation or facility: NMFS - Southeast Fisheries Science Center
   4. E-mail address: matthew.d.campbell@noaa.gov
   5. Phone number: 228-549-1690
2. **Responsible Party for Data Management***Program Managers, or their designee, shall be responsible for assuring the proper management of the data produced.*
   1. Name: Matthew D. Campbell
   2. Position Title: Gulf and Caribbean Reef Fish Branch Chief - Southeast Fisheries Science Center
   3. Name of current Position holder: New activity. We will detail individual plans for selected transition plans over the next month.