Data Management

The proposed data management plan for this project satisfies our need to manage large amounts of data, estimated at approximately 25 terabytes (TB). These data will include PIV flow measurements, high-speed holographic and standard videos, and time-resolved brightfield images.

Logistics of Data Management System

To host large quantities of metadata and results, The Gemmell lab has purchased a 96 TB Dell data storage system. The University of South Florida has a high-speed fiber optic internet allowing for high transfer capacity from 20 to 200 Mbps. We will also employ data collection on the recently developed database management developed by the Information Management Systems and Services group at USF. The IMSS database is cloud-based, enabling data sharing with collaborators and, eventually, anyone with web access.

Data Management Policy

<u>Data Description</u> – Data will include key experimental recordings, analysis results, and documentations. The key data include large amounts of high resolution, high speed videos, and PIV images, just to name a few. In addition, field experiments will produce high speed holographic movies and collimated brightfield image sequences. PI(s) will be responsible for defining the term "key". The analysis results are composed of experimental and simulation results, e.g. measurements of swimming, flow fields, and modeling results. Documentations include dissemination materials, publications, reports, procedures and protocols.

<u>Responsibility</u> – Data Manager, Brad Gemmell (USF), has overall responsibility for data management over the course of the research project and will monitor compliance with the plan in cooperation with key project personnel. The data manager working with IT technicians at USF oversees the deposit, collection, quality control and data submission, and dissemination and access control of these data.

<u>Data Archiving</u> – The research data from this project will be deposited with the digital repository at USF and also uploaded to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) repository. The ***frequency of data upload will be every 6 months*** over the duration of the project.

<u>Access and Sharing</u> – PIs will make the research ***data from this project available to the broader biology research community through regular uploads to the BCO-DMO database***. The severs are located at the Woods Hole Oceanographic Institution, and is a combination of the formerly independent Data Management Offices formed in support of the US JGOFS and US GLOBEC programs.

Data will also be made publicly available through access via a USF cloud-based database. *Public-use data files*: These files, including selected raw data and results after publications, can be accessed directly through the cloud-based database. *Term-use data files*: These file, including all raw data, intermediate results, and simulation results, can be accessed through the collaborating institutions. After agreeing to Terms of Use, the data can be accessed through each PI who is responsible for data. *Restricted-use data files:* These files include all data deposited by the owner (member of the collaboration) and users in the member institution with written permission granted by the owner. The data is accessed through remote access client to provide security and access controls. *Timelines*: The research data from the USF cloud-based database will be supplied to users determined by PIs. Users with restricted-use privileges can access the data immediately. Delayed access may be possible until the agreement is reached.

<u>Metadata</u> – Substantive metadata will be provided in compliance with the most relevant standard for biological science. New metadata standards have been developed in prior collaborations for data such as 3D holographic recordings where raw images are provided as well as links to reconstruction software. We will continue to be develop standards for modeling results and PIV images.

<u>Intellectual Property Rights</u> – Principal investigators and their institution hold the copyright for the research data they generate. Depositing data does not transfer the copyright.

<u>Storage and Backup</u> – USF IT will provide weekly backup services.