Data Management Plan

Types of data

1. Raw data: Nearly all experiments under the proposal involve electrophysiological recordings in behaving rats. The primary (i.e., raw) forms of data are: 1) waveform recordings in the format of the .plx files (Plexon MAT recording system); 2) video and LED tracking files in the form of .avt and .dvt files (the latter is output of the animal's head position in x,y coordinates collected at 60 Hz by the Plexon Cineplex Studio program); 3) histological data in the form of Nissl-stained brain slices (to localize recording or microinjection sites). The applicant's laboratory houses a microscope with wide-field lenses and a digital camera and regularly practices archiving of relevant digital photographs of brain slices..

2. Pre-processed data: Nearly all experiments will demand categorization of spike waveform data into subsets corresponding to single neurons. In the field, the method and outcome of such categorization is often a matter of interest. Waveform categorization output takes the form of .plx files (Plexon Inc.). Nearly all experiments involving navigation and tracking of the animal are put through a behavioral 'screening' procedure whereby portions of the behavioral record are selected for further analysis (e.g., identification of uninterrupted runs along a track). This process is carried out in Matlab and the interpretation of the output .mat files requires only a key to the behavioral coding scheme (easily available via the Nitz laboratory).

3. Analyzed data: All electrophysiological and behavioral data is analyzed using the commercially available program Matlab. The resulting .mat files involve analyses of many types. However, nearly all experiments involve creation of a 'ratemapping' analysis wherein the firing rates of neurons according to positions along a path are determined.

Data and Metadata standards

Electrophysiological data take the form of .plx files (Plexon Inc. standard), .avt files (overhead videos of recording arenas), .dvt files (Plexon Inc. standard for position-tracking data), or .jpg files (digital photographs of stained histological brain slices). Pre-processed data (see section above for explanation) comes in the form of .plx files (Plexon Inc. standard for categorized spike waveforms) and .mat files (Matlab standard – output of screening/filtering of tracking data). Analyzed data (see above section for explanation) takes the form of Matlab .mat files. Metadata comes primarily in the form of surgery and recording logs where the nature of the particular experiment (stereotaxic targets, channel configurations, behavioral apparati utilized, recording and animal nos., etc.) is written and drawn on paper. Such recording logs are copied once a month and stored offsite. Access could be made via digital photography if requested. Other metadata include the specific custom-written analysis programs utilized and information concerning amplifier settings and channel configurations. The latter are contained within the .plx electrophysiological data files. Custom-written analysis programs (Matlab .m files) are regularly copied and saved to several different on-site and off-site computers.

Policies for access and sharing

All data, metadata, and analyses collected under the proposed experiments will be made publicly available as per NSF guidelines within 2 years of collection via published manuscripts, publicly available final reports to NSF, and/or from data archives at UCSD's Department of Cognitive Science. The PI will take guidance from NSF concerning the right to use the data prior to opening it up to wider use. There are no ethical or privacy issues involved in sharing of this type of data and it is unlikely that sharing will incur more than modest cost. The most likely data-sharing scenario concerns the categorization of spike waveforms as this process does not enjoy agreement in standard over the entire field. Possible requests

may also come from neurophysiologists desiring to double-check the validity of a published analysis using categorized spike waveform data and behavioral tracking data.

Each of the formats by which data and metadata are archived can be used with relative ease by anyone requesting access. Plexon .plx files can be converted for use in standard programming environments such as Matlab through use of software downloaded from the Plexon website. Tracking data comes in the form of .dvt files which can be directly loaded into Matlab or Notepad. Video and digital photographic files are in formats (.avt, .jpg) that can be opened using a large number of freely available programs.

Data archiving

All electronic data are saved in triplicate with two copies kept on-site (UCSD Cognitive Science Department) and another kept off-site. As described, all data, whether in electronic or paper form, are copied, organized by animal number and recording day, and archived both on-site and off-site. There is no plan to destroy any collected data as the archive is not burdensome in cost or space. As such, data archives can be expected to be available for at least a period of ten years subsequent to publication of the relevant findings. As described above, the PI's attention to detail in archiving and the ease with which data can be shared in usable format makes data-sharing a relatively simple process for the proposed experiments. In the event that a simple email or snailmail delivery of electronic files cannot be accomplished, the PI has access to the department's system administrator who can set up an FTP server. A third avenue for sharing is via the SRB (Storage Resource Broker)/iRODS (integrated rule-based data) system developed at the San Diego Computer Center for sharing of data from NSF projects such as the Temporal Dynamics of Learning Center.