Abstract: DATA PROCESSING SYSTEM FOR GENERATING PREDICTIONS OF COGNITIVE OUTCOME IN PATIENTS, Y1

For over 60 years, intraoperative electrical stimulation in awake patients has been the gold standard for mapping functions in the human brain. During awake mapping, the patient is alert and comfortably positioned on the operating table. The goal of this clinical procedure is to identify where stimulation disrupts the patient's ability to carry out a given task. Regions of the brain critical for language ability can be localized by testing for disruption of a patient's ability to name pictures, read words, or repeat sentences. The current project consists of a software application that will be used to gather data, and perform spatial and temporal co-registration of EEG, fMRI, brain-mapping and behavioral data. The solution also has a client application (on-site) that will be used to present different types of stimuli before-during and after awake craniotomies on patients that have different brain lesions with the main objective to reduce the likelihood to cause any deterioration of any critical cognitive ability. The whole system consists of a client application that gathers all the different brain imaging data, then it anonymizes all the relevant information (taking into consideration HIPAA and FDA regulations). Next, all the anonymized data gathered gets stored into a data repository that feeds the co-registration, and normalization system that will output a visualization model. This process will be enriched by the Predictive Analytics model that we are trying to build by using different Machine Learning/Deep Learning/ Statistical models to minimize the risk associated with brain tissue removal.

With respect to the Regular Memory, we request 10,000 units since we are planning to create an environment that consists of at least 4 virtual machines. From those 4 VMs, 2 of them will be dedicated to support the data management system.
With respect to the storage requested, our lab has historical records that will be used to train the initial ML model. For now, we have data from 50 patients. The data consists of scans from a wide variety of medical scans such as MRI, fMRI, DWI, EEG, ECoG, etc. Also, in some cases, we have medical and multimedia records that enriches the dataset. The amount of information varies from patient to patient, but the average is 5GB. We estimate that once this project starts collecting more information from different medical centers, the amount of data could be doubled within a couple of months.