***Data Management Plan***

**Expected Data.** We employ state-of-the-art control systems for accelerator operation with fully automated computer-controlled data recording techniques, so that all measurements will yield raw data in electronic file format. Data are recorded using a variety of software packages, some commercially produced and some specific to the accelerator facilities. Raw data are sometimes processed utilizing standard software packages (MATLAB, etc.) or custom visualization or analysis code. Raw and processed data are used as inputs to models and computer simulations, whose algorithms, computer codes, and outputs are parts of complete datasets. Details of the simulations and models will be published in peer-reviewed journals, and documented code and algorithms will be included as part of the complete dataset.

In particular, all storage ring parameters are routinely recorded by the CESR control system. Data collected with beam instrumentation, including turn by turn and bunch by bunch beam position and bunch size monitors, and electron cloud detectors are archived, as are all relevant calibration constants. For a subset of measured quantities a “pass 1” processed version of the data are stored. We continue to add storage capability in order to maintain complete records

**Data Format.** Hardcopy notebooks will be maintained to document installation and use of project equipment and experimental procedures developed over the course of this project.

The data themselves will be stored in ASCII, CSV, or binary formats (sometimes using typical MATLAB I/O standards and functions) for large size sets. Data snapshots will be stored in individual directories with appropriate time stamps. Indexing of the experimental data (metadata) will be available in each directory and the main parent directory, along with documentation in the hardcopy notebooks according to accepted standards in the lab. Metadata standards will be developed by the investigators as necessary (custom) to unambiguously associate data with specimens and preparation procedures, instruments, dates and times, processing and analysis techniques, and any other information deemed necessary to fully understand the data analysis flow and results.

**Data Archival.** All raw, processed and analyzed data will be copied to a storage area network of redundant disk arrays with error protection and a total capacity of hundreds of terabytes. Each folder or directory on this file system will include data index tables and summaries of experimental data contained therein. The data on these disk arrays and all of its content are archived on a revision controlled redundant-hardware off-site tape library with daily incremental and full monthly backups. The hardcopy notebooks will be stored in file cabinets along with the equipment.

To guarantee redundant access and persistence, the PI will also archive all project data in the eCommons@Cornell repository. This Digital Repository is powered by opensource project DSpace and is open to anyone affiliated with Cornell University as a place to capture, store, index, preserve and redistribute materials in digital formats that may be useful for educational, scholarly, research or historical purposes. eCommons@Cornell is a service of the Cornell University Library that provides long-term, open access to a broad range of Cornell-related digital content of enduring value. All work deposited in eCommons will be assigned a persistent identifier and a persistent Web address (URL), guaranteeing that data will be accessible independent of the hosting location or domain of the repository.

**Access to Data and Data Sharing Practices and Policies.** We expect that all the data gathered and results derived during this project will be freely available to interested parties. A number of collaborative Wiki-based websites are maintained by the Laboratory for research areas covering CesrTA, These web-sites contain many details of specific hardware design and instrumentation, weekly meeting minutes, summary of experimental procedures, as well as preliminary experimental results. The access to these resources is password-protected with our collaborators outside of Cornell having full access to these resources and contributing frequently to its content.

**Policies for Re-Use, Re-distribution.** All of our websites will include standard disclaimers and conditions regarding the use of the data in other publications and products in accordance with Cornell University standards. In summary, data is allowed to be reused without expressed consent from the PI’s but requires statements of attribution and disclaimers that the originators of the data are not responsible in any way for re-use or novel interpretations or results. Disclaimers and policies for deposit, access, withdrawal, ownership, privacy and preservation of data will follow closely those for the eCommons@Cornell depository found at <http://ecommons.library.cornell.edu/policy.html>.

Machine operations and people data (proposal, beamtime requests, etc.) are always kept private and secure (no external access to raw data). Machine and facility use information is summarized periodically for official reporting purposes only (i.e. NSF review committees, GPRA and NSF annual reports).