ICS Language for Inclusion in Data Management Plans

The following is language for PIs planning to use the Institute for CyberScience’s Advanced Cyber Infrastructure (ICS-ACI) for their data storage and preservation purposes. The language can be used in PIs’ Data Management Plans (DMP). Given the typical brevity of DMPs (NSF DMPs are two pages or shorter), the description gives a broad and brief sense of ICS-ACI’s data storage, preservation, and security capabilities. If using the language below in your proposal, please read carefully to ensure that all of the claims meet your specific use case.

**Data Storage and Preservation**

Over the course of the research project, research data will be hosted by the Pennsylvania State University’s Institute for CyberScience (ICS) through its Advanced Cyber Infrastructure (ICS-ACI). ICS-ACI provides both active storage (for data that is being worked on, requiring frequent access) and near-line storage (for back-up purposes and data that needs only infrequent access). Active storage is achieved through DDN 12KX40 and GS7K flash storage array systems, while near-line storage utilizes Oracle’s FS1 flash storage appliance and a SL8500 Tape Library. Active storage is backed up to the SL8500 daily.

We will also use ICS-ACI to archive the research data for at least three years after the end of the award or after public release of the data, whichever comes later. ICS provides long-term archival services using the Oracle SL8500. Multiple copies of archived data are created through Oracle Hierarchical Storage Manager (Oracle HSM) to safeguard against data corruption. Oracle HSM generates and maintains metadata on archived files so that the data can be readily accessed. This technology affords us easy retrieval of data, even years after it has been written.

**Data Security**

ICS implements various security measures to ensure that data stored on the ICS-ACI system remains safe. ICS-ACI requires a strong password and two-factor authentication for access, and all access can be audited by ICS staff. To mitigate the potential for malicious software and security attacks, ICS-ACI employs automated weekly scans for identifying and patching software vulnerabilities. ICS-ACI provides the capability to encrypt data in-flight (when moving between points) and at rest (while written in storage). ICS-ACI login/endpoint nodes are protected by software-based firewalls that only permit Secure Shell (SSH) traffic. By default, ICS enforces “Least Privilege” access concepts across the system, providing users with only the minimum set of permissions and accesses required to complete their function.

ICS-ACI storage is physically protected in Penn State’s Tower Road Data Center (TRDC). Physical access to the systems is limited to systems administration personnel with exceptions controlled by the TRDC’s secure operations center. TRDC requires swipe-card access and is monitored at all times.

Data stored on ICS-ACI’s active storage systems is backed up to tape storage for a period of 90 days. Backup data is automatically purged from tape once the 90 days has been exceeded.