G Monarch II: Cyber Corps Through Transformation - Renewal: Data Management Plan

This section describes how the data generated from the effort described in the proposal will be collected, stored, and made available. Our project data is of two major types:

- data associated with participant selection and participation in the proposed program and
- data generated as part of thesis research conducted by participants.

Because student thesis research topics cannot be known \textit{a priori}, the overall data management plan for theses is described. Any issues not specifically referenced in this project’s Data Management Plan will fall under the overarching Naval Postgraduate School Data Management Policy.

G.1 Types of Data

G.1.1 Student Information

Student information maintained by NPS is not intended for research, but exclusively for: admissions, tracking academic progress, and to assist students with internship and job searches. Institutional Review Board (IRB) approval is not required for administrative use of student information. Student information falls into two categories: Personally Identifiable Information (PII) and information that may be public.

Student PII includes:

- Application packages, which include transcripts, GRE scores, recommendation letters, etc.
- Copies of NPS transcripts, or other grades
- Private information regarding internships and jobs. Sometimes students discuss offer letters with us. Although we do not make copies of these, the information must be protected

Student application packages are held by the NPS Office of Admissions and for our program-internal admissions screening processing.

An example of student information that may be made public is a resume to be passed to potential employers.

G.1.2 Theses and papers

The primary data products of this effort will be the Masters Theses of the student participants. Scholarly papers and technical reports are also anticipated products of this effort.

G.1.3 Other Artifacts

Source code, data, document source files, etc. may be produced during thesis research.

G.2 Data and Metadata Standards

Student data will be encoded using industry-accepted formats for text documents, spreadsheets and databases. Data collected as part of student thesis research will be encoded in appropriate, generally accepted formats. All data formats will be clearly indicated, by using file types or more detailed data descriptions.

G.3 Data Access and Information Security

All PII is protected. The NPS Office of Admissions maintains strict procedures for the PII protection. The NPS SFS program also carefully protects student PII. Locked filing cabinets are used to store paper documents. Electronic information is password protected and available only to those with a need to know. Expired information, such as applications of individuals not admitted to the program is destroyed either through deletion or shredding, unless a student requests that it be retained for the next admissions cycle.

Student resumes may contain PII such as an email addresses and phone number. Resumes are only sent to others, such as a potential internship host or employer, with the student’s permission.
Reports regarding the Scholarship for Service program may contain statistics regarding the demographics of the applicants and of the students matriculated into the program. This information will be filtered and sanitized so that individual students cannot be identified.

Data used in thesis research will be protected in accordance with the sensitivity or classification of the information. This includes data that may be produced under some form of non-disclosure agreement, as might be in place under a Cooperative Research and Development Agreement (CRADA). Mechanisms for data protection during and following research may include the use of VPNs, encryption, safes, SECRET-level facilities, and storage in the NPS Sensitive Compartmented Information Facility (SCIF).

G.4 Information Redistribution

Redistribution of publicly available information will be through the archival vehicles described in Section G.5.2. Unclassified theses, technical report, and scholarly papers will be publicly available. Redistribution of information gathered or produced during student thesis research will be at the discretion of the student’s thesis advisor. It is expected that the redistribution vehicles will among those described in Section G.5.2. Requests for sensitive or classified information will be handled on an individual basis. In the case of material produced under non-disclosure or Cooperative Research and Development Agreements (CRADAs), access may be denied. Intermediate results will not be made available.

G.5 Data Preservation

G.5.1 Project-Internal Student Research Data Preservation

Source code, data, etc. produced during student thesis research are internally archived using a check-in process. An example is given here. We use the subversion configuration management tool. Prior to graduation, students must perform an archive task. The main steps are: (1) Engineering validation of the final implementation. This includes verification of the installation procedures and testing procedures, as required by the particular project. (2) Code submission for the Configuration Management archive.

Archive Task: After a possible Engineering Validation Task is completed, the student must prepare a configuration management (CM) submission package containing the following:

- Archive materials, i.e., CD-ROM and tapes, created during pre-check-in engineering validation
- Hard-copy of the README file on the CD-ROM

The student must provide the submission package to the thesis advisor, who forwards it to our engineering staff for CM submission.

G.5.2 General Archival Preservation of Information

The NPS and the Department of Defense provide for archival storage and indexing of student theses. All unclassified and SECRET theses are stored electronically in the NPS library. Access to the former is through the library’s online catalog. Access to the latter is limited to authorized individuals. More sensitive theses are stored in the NPS SCIF. Unclassified theses are sent to the Defense Technical Information Center (DTIC), where they are also indexed, archived, and made publicly available via the DTIC website.

Naval Postgraduate School technical reports are archived both in the Dudley Knox Library and by DTIC. Journal articles, conference papers, and workshop papers are archived by the publishing body, such as the ACM and the IEEE. Both organizations maintain on-line libraries. Investigators also maintain websites from which their publications may be downloaded.

In addition to its standard library inventory used for theses, the NPS library provides a facility, Calhoun, for archiving a wide range of digital artifacts including copies of theses, scholarly publications, technical reports, and other digital artifacts. Works in Calhoun are visible to search engines and to metadata harvesters such as OpenDOAR. Calhoun adheres to the concepts of the Open Archive Initiative, which facilitates access via a global network of interoperable repositories.