Plan Overview

Title: Minority Cyber-CREWS: Minority Community Cybersecurity Research, Educational Engagement, Workforce Development, and Security Best Practices

Creator: Gahangir Hossain

Affiliation: University of North Texas (unt.edu)

Principal Investigator: Gahangir Hossain

Data Manager: Gahangir Hossain

Funder: Department of Homeland Security (DHS)

Funding opportunity number: DHS-23-ST-SLA-001

Grant: 23STSLA00013-01-00

Template:

Project abstract:

This minority Cyber-CREW proposal seeks funding from the Department of Homeland Security (DHS) S&T Office of University Program for FY2023 Scientific Leadership Awards to improve cybersecurity research, education, and collaboration capabilities for the University of North Texas (UNT), a minority-serving institution (MSI) in the rapidly growing and diverse Dallas-Fort Worth (DFW) metropolitan area. The project's funding will be used for cybersecurity research and development for underrepresented undergraduates and early-career faculty members, as well as the creation of cybersecurity education pathways for community colleges and certificate programs to enhance cybersecurity skills. The project will also provide summer research opportunities for undergraduates in areas such as cyber-ethics, privacy, trust, and fraud investigation, blockchain, cyber datamining and machine learning, and web application and game development. Faculty, including early career junior members, will conduct rigorous research on cyber data analytics, machine learning (ML), artificial intelligence (AI), malware analysis, software assurance, and mobility with critical cyberinfrastructures including energy, agriculture, and healthcare services with an emphasis on how the research impacts and/or empowers minority communities. The research team will develop novel software tools with algorithms, prototypes, models, and maintenance as well as strategies for advancing diversity, equity, inclusion, and belonging within the DHS cybersecurity initiatives. The team will also design modern cybersecurity courses and certificate programs for all STEM students in the Department of Information Science. The research team will create a pathway for community college students to transfer to the four-year program at UNT, providing opportunities for re-skilling or up-skilling towards cybersecurity careers after completing an associate degree at their community college.

Last modified: 07-12-2023

Copyright information:

Minority Cyber-CREWS: Minority Community Cybersecurity Research, Educational Engagement, Workforce Development, and Security Best Practices

Types of data

The Data Management Plan should describe the types of data, metadata, scripts used to generate the data or metadata, experimental results, samples, physical collections, software, curriculum materials, or other materials to be produced in the course of the project.

The data used for this project will include students' names, address, and their educational information. This information will be coded with their ID. Some of this data will be restructured and analyzed to produce new data sets, visualizations, and reports.

Data and metadata standards

The Data Management Plan should address the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies). It should also cover any other types of information that would be maintained and shared regarding data, e.g. the means by which it was generated, detailed analytical and procedural information required to reproduce experimental results, and other metadata.

The project will work together with the University of North Texas Digital Libraries team to select file formats that ensure optimal and secure long-term management of digital data and information. The choices of file formats for text and images will be guided by best practices in data curation and digital preservation. When utilizing proprietary software, we will make every possible effort to generate output in open formats.

To provide comprehensive information about each digital resource in the project, metadata records will be generated. These records will be preserved and made openly accessible through the UNT Digital Library.

Policies for access, sharing, and privacy

The Data Management Plan should address the policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. It should cover any factors that limit the ability to manage and share data, e.g. legal and ethical restrictions on access to human subject data.

The project team will collaborate with UNT Libraries to assess the information objects and derived data, evaluating their potential benefits for future researchers. Accordingly, these valuable resources will be made publicly accessible through the UNT Libraries' Digital Library for long-term preservation. Throughout this process, the team will prioritize the protection of privacy, confidentiality, and any proprietary or intellectual property rights. For data sources that are subject to embargo for a certain period, metadata records will be available to facilitate the discovery of these resources, even if the data itself is not immediately accessible. All presentations and publications related to the project will be accessible through UNT's open access Scholarly Works Repository, which is hosted in UNT's Digital Libraries. You can find them at http://digital.library.unt.edu/explore/collections/UNTSW/.

Policies for re-use, re-distribution, derivatives

The Data Management Plan should address the policies and provision for re-use, re-distribution, and the production of derivatives.

The project aims to make the data and derived datasets collected and/or created as widely available as possible for reuse by other researchers. The availability of the data will be determined in consultation with all stakeholders, based on the appraisal process mentioned earlier. If the project team determines that certain data or datasets require limited access, researchers can submit requests to the Principal Investigator. Metadata associated with all data and derived datasets will provide information on how to request access and use the data. When utilizing the project's collected, created, and/or derived data products, users will be kindly requested to provide a standard citation that acknowledges both the Department of Homeland Security (DHS) and the project itself. This proper attribution will ensure due credit is given to both entities for the access and utilization of the data.

Plans for archiving and preservation.

The Data Management Plan should address the plans for archiving data, samples, and other

research products, and for the preservation of access to them. It should cover the period of time the data will be retained and shared; how data are to be managed, maintained, and disseminated; and mechanisms and formats for storing data and making them accessible to others, which may include third party facilities and repositories.

The UNT Digital Library is dedicated to ensuring the long-term availability and management of publicly accessible research outputs. Recently, the UNT Libraries conducted a self-audit based on the Trusted Repositories Audit & Certification: Criteria and Checklist (TRAC). You can find more information about the audit at https://library.unt.edu/digital-libraries/trusted-digital-repository/.

To safeguard the data generated by this project, the UNT Data Repository (http://digital.library.unt.edu/datarepository) will be utilized. The data will be regularly backed up to protect against potential loss resulting from hardware failures, fire, theft, and other unforeseen circumstances. All research products and project-related materials, including technical reports, presentations, and publications, will be made accessible for the long term. They will be hosted in the UNT Scholarly Works open access repository (http://digital.library.unt.edu/scholarlyworks), which is part of the UNT Digital Library.

Roles and responsibilities

The Data Management Plan should clearly articulate how the PI and co-PIs plan to manage and disseminate data generated by the project. The plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data, and consider changes that would occur should a PI or co-PI leave the institution or project. It should describe how the research team plans to deposit data into any relevant and appropriate disciplinary repositories that are appropriately managed and that are likely to maintain the metadata necessary for future use and discovery. Any costs associated with implementing the DMP should be explained in the Budget Justification.

The dissemination of the results generated by this project will be conducted through various means. We are seeking funding for publications to support travel expenses for the Principal Investigator (PI), co-PIs, and graduate research assistant. Major conferences in the field that we plan to target include:
- ACM/IEEE-CS Joint Conference on Digital Library (JCDL): https://www.jcdl.org/
- Digital Libraries Forum (DLF): https://www.diglib.org/dlf-events/
- International Conference on Knowledge Management (ICKM): www.ickm.net/
- Chicago Forum on Global Cities: https://www.thechicagocouncil.org/
Given the complex and collaborative nature of this project, there is potential for further study across various aspects of creative data design, effective technology application, and community practices. These explorations could lead to additional data production that warrants publication in peer-reviewed journals. A few notable journals in this domain include:

- International Journal of Business Intelligence and Data Mining:
https://www.inderscience.com/jhome.php?jcode=ijbidm
- Journal of the American Society for Information Science and Technology:
https://onlinelibrary.wiley.com/journal/15322890
- Big Data: https://home.liebertpub.com/publications/big-data/611/overview
- Big Data Research: https://www.journals.elsevier.com/big-data-research/

These journals provide platforms for sharing research findings and engaging with the scholarly community in the respective fields.