T^3-CIDERS: Fostering a Community of Practice in CI- and Data-Enabled Cybersecurity Research Through A Train-the-Trainer Program

Wirawan Purwanto
Old Dominion University
Norfolk, Virginia
wpurwant@odu.edu

Mohan Yang
Texas A&M University
College Station, Texas
mohanyang@tamu.edu

Peng Jiang
University of Nebraska Omaha
Omaha, Nebraska
pjiang@unomaha.edu

Masha Sosonkina
Old Dominion University
Norfolk, Virginia
msosonki@odu.edu

Hongyi Wu
The University of Arizona
Tucson, Arizona
mhwu@arizona.edu

ABSTRACT

We present a training program named T^3-CIDERS, the "Train-The-Trainer approach to fostering cyberinfrastructure (CI)- and Data-Enabled Research in CyberSecurity." T^3-CIDERS is a train-the-trainer program for advanced cyberinfrastructure (CI) skills that is designed to be synergistic with research, teaching, and learning activities in cybersecurity and cyber-related disciplines. The participants, termed "future trainers" (FTs), are trained in effective instructional design and CI hands-on materials from "DeapSECURE", developed in a previous CyberTraining program. T^3-CIDERS aims to enhance cybersecurity research and education through broader adoption of advanced CI techniques such as artificial intelligence, big data, parallel programming, and platforms like high-performance computing (HPC) systems. T^3-CIDERS includes pre-training, a weeklong summer institute, ongoing learning engagements, and local training activities. The FTs conduct local training tailored to the needs at their respective home institutions. They will also develop a new CI training module (called "Module X") based on the observed common needs in the cybersecurity research community. Community building is integral to T^3-CIDERS as its overarching goal. The first cohort of FTs who took the 2024 summer institute comprises faculty members, researchers, and students representing multiple states.

KEYWORDS

Cyberinfrastructure, cybersecurity, research, train-the-trainer, pedagogy, HPC, parallel computing, big data, machine learning

1 INTRODUCTION

Cyber-enabled technologies have revolutionized the 21st century society by increasing efficiency, convenience, and productivity in every walk of life. With the increasing volume and sophistication of cyberattacks, researchers, engineers, and practitioners heavily rely on advanced cyberinfrastructure (CI) techniques to assess cyber risks, identify and mitigate threats, and achieve defense in depth. CI capabilities such as artificial intelligence (AI), machine learning, big data, as well as advanced CI platforms, e.g., the cloud and high-performance computing (HPC) have become integral to support the state-of-the-art cybersecurity research and implementations.

Despite the widespread availability of CI and its importance in strengthening the security of cyber systems, there is still a significant gap in the preparedness of the cybersecurity workforce to leverage CI. In research, the lack of preparedness in CI often hampers, or frustrates, students from entering research areas that rely heavily on CI, such as AI/cybersecurity intersection and privacy-preserving machine learning techniques. To fill this gap and broaden CI adoption in research and education in cybersecurity, we developed the "Train-The-Trainer Approach to Fostering Cyberinfrastructure (CI)- and Data-Enabled Research in CyberSecurity" (T^3-CIDERS) training program [9]. The goal of T^3-CIDERS is to accelerate state-of-the-art research and development in cybersecurity and related fields by (1) preparing competent trainers to broaden the utilization of advanced CI in these disciplines; and (2) fostering a CI-enabled cybersecurity research community of practice. As a train-the-trainer program for advanced CI, T^3-CIDERS is designed to be synergistic with research, teaching, and learning activities of its participants.

The T^3-CIDERS program is a year-long engagement for each cohort of participants, referred to as *Future Trainers* (FTs). The FTs are chosen from faculty, researchers, students, and practitioners with interests in cybersecurity and cyber-related fields. They will receive training on the basic skills in key CI areas and effective pedagogical methods to further promulgate these CI skills to others in their academic communities. The FTs are clustered into units ("FT units"), each of which consists of one faculty or researcher and one or two students. As part of the requirements of the program, these FT units will collaboratively prepare and conduct CI training activities at their respective institutions. The activities of T^3-CIDERS encourage collaborations and community building among FTs as a CI- and cybersecurity-focused community of practice.

T^3-CIDERS leverages and builds upon "DeapSECURE" introductory CI training modules [5–8], which covers fundamental CI topics: HPC, Big Data, Machine Learning, Deep Learning, Cryptography,

https://doi.org/10.22369/issn.2153-4136/x/x/x

and Parallel Programming. As a minimum prerequisite to participate in T^3-CIDERS, FTs need to have working knowledge of computer programming (in any language); preferably, they should also have experience working with one or more of the aforementioned CI topics.

2 OVERVIEW OF THE T^3-CIDERS PROGRAM

2.1 Program Structure

T^3-CIDERS is designed to train and engage the FTs through be a year-long program. This comprehensive approach consists of six major phases:

Phase 1: Pre-training — A month-long, self-paced, virtual training on CI skills baseline, covering HPC, Unix shell, Jupyter, and Pandas. This phase ensures that all participants have a foundational understanding of essential CI concepts and tools.

Phase 2: Summer Institute — An intensive, week-long, in-person training that covers an overview of CI techniques (Machine Learning, Deep Learning, Cryptography, and Parallel Programming), as well as pedagogy/training methodology which includes educational foundations, assessment, hands-on lesson planning, and simulated teaching practice. It also includes discussion panels with experts aimed to broaden the FTs' perspective on pressing issues in cybersecurity practice and education.

Phase 3: Monthly Learning Engagement — After the summer institute, the FTs will engage in year-long virtual meetings (about once a month) to further their CI learning journey, share experiences in CI-augmented teaching and/or research, connect with peers and cyber experts. This ongoing engagement fosters an inclusive and supportive community of practice and encourages continuous professional development.

Phase 4: Local Training — As part of their enrollment with T^3-CIDERS, FT units are required to conduct local training activities that are tailored to their academic communities, ensuring the dissemination of CI knowledge and skills.

Phase 5: K-12 Outreach — FT units have the opportunity to spread awareness of cybersecurity, research in cybersecurity, and the role of CI to support cyber research in local K-12 schools, inspiring students to be the next generation of cybersecurity professionals. This phase is optional, but FTs are strongly encouraged to reach out to the communities around them in this way.

Phase 6: "Module X" — A community-driven initiative where FTs and the project team will collaboratively design and develop new CI training module(s) to address additional training gaps in the preparation of students for cybersecurity research.

The pre-training and summer institute serve as a short-term "boot camp" to the FTs. The monthly learning engagement is an effort to foster the connections among FTs that were established during the summer institute and provide continuous learning opportunities. The local training activities, K-12 outreach, and "Module X" represent outcomes and sustained activities facilitated by the T^3-CIDERS program.

2.2 CI Skills Baseline

T^3-CIDERS equips FTs with understanding of baseline CI skills that can be further taught to their own learners, who most likely are novice to CI. The baseline skills are provided by adapting the DeapSECURE's lesson modules[6] covering HPC, Big Data, Machine Learning, Deep Learning, Cryptography, and Parallel Programming. DeapSECURE lessons emphasizes hands-on introduction to these topics and incorporates research datasets and use cases to provide strong relevance to cybersecurity. Table 1 shows the six lesson modules, the key contents of each lesson, the software tools introduced, and the target cybersecurity application used throughout the lesson. DeapSECURE modules integrate the principle of authentic learning[3]: Certain challenges (problems), carefully selected within the target applications, motivate the introduction of the CI methods and/or techniques. The learners are guided throughout a series of hands-on activities (similar to Carpentries-style live coding method[4]) to leverage these methods and address the posed challenges. In this way, CI topics are introduced in a way that is relevant to cybersecurity applications.

The DeapSECURE lessons were originally designed for non-degree, non-credit, bootcamp-style training targeting novice learners who may not know the CI topics at all. The lessons provide a gentle onramp by covering only the most critical concepts and techniques that novices must learn to become competent learners and practitioners on their own. The lesson materials include Carpentries-style online e-textbooks, Jupyter notebooks, and hands-on files (sample scripts and templates, datasets, etc.). as explained further below. In addition to providing baseline CI skills, these lessons also serves as teaching resources, from which the FTs can draw and/or adapt to suit their own learners' needs.

2.3 Training Infrastructure

The T^3-CIDERS program leverages the following platforms as its infrastructure: (1) ODU's "Wahab" HPC cluster for the hands-on learning environment; (2) Canvas learning management system to present the lesson materials to the FTs; (3) Discord instant messaging platform, for discussions among FTs and with the training team; (4) LinkedIn and X (Twitter) as the social media for public outreach and engagement; (5) Qualtrics for data collection and surveys; (6) Gitlab for collaborative hands-on lesson development.

3 THE 2024 COHORT OF FUTURE TRAINERS

The 2024 summer institute took place in July 29–August 2 at the Norfolk campus of Old Dominion University. It drew together 10 faculty/researchers and 12 students with varied backgrounds. These FTs represented multiple states (Washington, Texas, Virginia, and New York), reflecting the program's broad geographic reach.

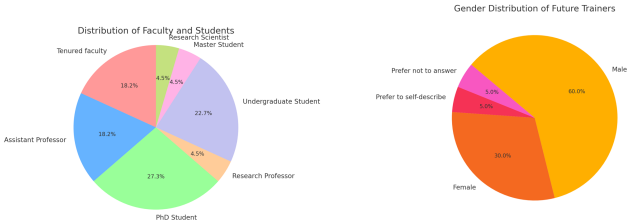
3.1 Participant Background

Figure 1 highlights the distribution of participants by role and background. As shown in Figure 1a, 36% of the total participants were tenured or assistant professors, while 9% held positions as research scientists or professors. Among the student participants, 31% were graduate students, and 23% were undergraduate researchers who were attracted to the program for its academic and professional opportunities. As displayed in Figure 1b, 60% of participants identified as male, 30% as female, 10% as others.

The major distribution of participants (Figure 1c) illustrates strong alignment of our proposed program to the research interests of our participants. Of the 22 participants, 8 are majoring in

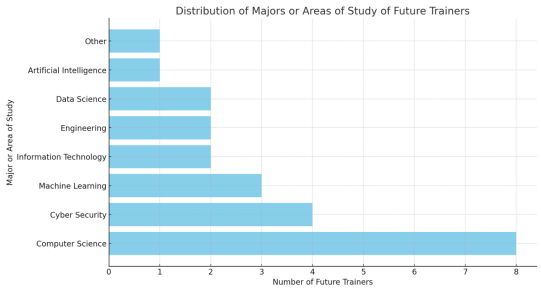
Table 1: Currently available DeapSECURE's CI lesson modules

Module Name	Description	Toolkits	Target Application
Introduction to HPC	Intro to HPC platform and basic parallel processing	UNIX shell commands, SLURM job scheduler	Analysis of a spam collection
Dealing with Big Data	Processing, cleaning, analyzing, and visualizing big data sets	Pandas, Matplotlib, Seaborn	Data preparation and smartphone app classification using Sherlock dataset
Machine Learning	Workflow of building, training and validating machine learning models	Scikit-learn	
Deep Learning using Neural Networks	Building, training and validating neural networks	KERAS, TensorFlow	
Cryptography for Privacy-Preserving Computation	Symmetric cryptography; homomorphic encryption for privacy-preserving computation	AES-Python [10], PyCryptodome [1], Python-Paillier [2]	Image encryption with Paillier crypto
Parallel & High-Performance Programming	Parallel programming with MPI	mpi4py	



(a) Faculty-student ratio

(b) Gender distribution



(c) Major distribution

Figure 1: The statistical overview of 2024 cohort

Computer Science, 4 in Cybersecurity, and another 4 in Artificial Intelligence/Machine Learning (AI/ML). Additionally, there are 2 participants each in Data Science, Engineering, and Information Technology. The mutual focus in CI- and Cybersecurity-related topics serves as the cornerstone in building a community of practice of next-generation cybersecurity researchers and scholars in leveraging CI technologies to accelerate state-of-the-art research and development in cybersecurity and cybersecurity-related fields.

3.2 Program Details

3.2.1 Pre-Training. Prior to the summer institute, FTs were onboarded to Canvas as the platform to host lesson materials, and to Discord (a group messaging platform used to facilitate discussions, ongoing collaborations, and community building). Each FT was required to complete the two pre-training modules, which helped them to get familiar with the HPC, as well as the prerequisite knowledge for the summer institute. Two virtual Q&A sessions were held to answer questions related to access to the ODU HPC cluster, and the technical challenges encountered in the pre-training modules.

3.2.2 Summer Institute. The Summer Institute for the 2024 Cohort was designed to provide an intensive, hands-on learning experience, combining technical lectures with practical activities, pedagogical training, and collaborative and interactive lesson design sessions. Participants engaged with well-designed cutting-edge topics while applying their knowledge through hands-on sessions and collaborative discussion. Additionally, pedagogical training is involved throughout the week, emphasizing the importance of effective teaching in the technical fields. Each FT unit has their own studio time in the afternoon to prepare a draft lesson plan for a local training activity that will be further developed and executed in the 2024–2025 academic year. On the last day of the summer institute, each FT unit presents a 15-20 teaching demo to showcase the lesson plans they built during the studio times, coupling with the pedagogical methods introduced in the summer institute.

3.2.3 Professional Developments. Throughout the week-long workshop, we organized two panel discussions and a field trip to bridge the gap between academic learning and real-world cybersecurity practices. The first panel featured industry partners from Hampton Roads, who provided a showcase of cutting-edge cybersecurity techniques and technologies. They discussed the challenges faced in real-world cybersecurity operations and highlighted advanced CI techniques being used to solve these problems. This session offered participants valuable insights into the practical application of their skills in industry settings. The second panel brought together high school teachers who had participated in the NSA GenCyber Summer Camps hosted at ODU in the past two years. These educators shared their experiences teaching cybersecurity concepts to their own students, offering a unique perspective on K-12 education. Our FTs engaged in productive discussions with the teachers, these exchanges provided our participants with a deeper understanding of how cybersecurity education is being approached at the K-12 level and the opportunities for integrating emerging technologies into the curriculum. The field trip to Old Dominion

University's ITS data center offered participants a firsthand look at the infrastructure supporting high-performance computing and cybersecurity efforts, allowing them to observe real-world applications of the concepts discussed throughout the workshop.

3.2.4 Monthly Engagement. After the conclusion of the summer institute and continuing through the end of each cohort year (April 2025), the project team will maintain regular engagement with the FTs through monthly virtual events. The initial engagements will primarily focus on lesson plan development. The project team will monitor progress from each FT unit on their lesson development and local training planning, providing feedback and support to help them refine their plans until the local training is implemented. From 2025, monthly research webinars will be hosted for both FT units and the broader community, allowing faculty to present their research findings. Additionally, we plan to invite external speakers to discuss recent advancements in HPC, education, cybersecurity research, etc, extending the influences to a wider community.

3.3 Assessment and Impacts

We plan to conduct a multiple-level and multiple-phased evaluation approach before, during, and after the program. The evaluation will include both formative and summative plans with a mix-method design for data triangulation. T^3-CIDERS employs multiple methods such as pre-survey, daily reflection and feedback form, post-institute survey, interviews, and final (post-) survey to assess the program's contents and effectiveness, FTs' attitude, learning, and teaching plan. (Since the cohort is still ongoing as of the time of writing, the evaluation has been conducted only partially.)

3.3.1 Before the summer institute. A pre-survey was sent out to the FTs and collected their feedback before the first day of the summer institute. It includes questions to identify FT's demographic information, characteristics, reactions toward the program, motivation, needs, and attitudes that need to be considered in designing and redesigning the program. It also provides insights into FT's target learners and helps us further tailor the content and instruction.

3.3.2 During the summer institute. During the summer institute, observational notes were taken to track the FTs' behavior, interactions with peers, and engagements. This helped in assessing participant involvement and learning application. FTs were encouraged to submit daily reflections, documenting their experiences and thoughts during the program. These sessions help ensure that learning is an iterative process, with participants and our project team, continually improving their understanding and teaching methods.

3.3.3 After the summer institute. Following the completion of the summer institute, a post-institute survey was launched to collect participants' reaction with the training they received thus far. Furthermore, after each FT unit conducts their local training or outreach, a survey will be distributed to their attendees to gather feedback and assess the effectiveness of their training efforts. At the end of the cohort period, a post-survey will be conducted to assess participants' satisfaction with the entire program, their views on the relevance of the content, and their confidence in applying the skills they acquired. These results will be compared with the pre-survey to identify any shifts in attitudes and technical abilities that occurred throughout the program. As each cohort year comes to a close, we will begin examining the Return on Expectations (ROE). This will involve conducting semi-structured interviews with participants, the program coordinator, and training mentors to gather insights into their experiences, the program's structure, activities, content, and instructors.

4 CONCLUSION

T^3-CIDERS is a promising initiative that addresses the growing need for advanced CI expertise in cybersecurity research and education. By providing comprehensive training, fostering a supportive community of practice, and promoting sustained engagement, the program is effectively preparing the next generation of cybersecurity professionals to tackle emerging challenges. The positive feedback from the 2024 cohort, the anticipated local training activities during the 2024–2025 academic years and beyond, as well as the ongoing development of "Module X" demonstrate the program's potential for long-term impact.

5 ACKNOWLEDGMENTS

Funding for this program comes from the U.S. National Science Foundation CyberTraining grants #2320998 and 2320999. The authors graciously acknowledge the support from ODU School of Cybersecurity for hosting the 2024 summer institute, and ODU Research and Cloud Computing group for the use of their Wahab HPC cluster for conducting the CI training activities. Dr. Nitin Sukhija from Slippery Rock University of Pennsylvania reviewed the pre-training materials.

REFERENCES

[1] Simon Arneaud, Nevins Bartolomeo, and Thorsten E. Behrens *et al.* [n. d.]. PyCryptodome. https://pycryptodome.org/

[2] CSIRO's Data61. 2013. Python Paillier Library. https://github.com/data61/python-paillier

[3] Jan Herrington and Ron Oliver. 2000. An Instructional Design Framework for Authentic Learning Environments. *Educational Technology Research and Development* 48, 3 (2000), 23–48. https://doi.org/10.1007/BF02319856

[4] Alexander Nederbragt, Rayna Michelle Harris, Alison Presmanes Hill, and Greg Wilson. 2020. Ten quick tips for teaching with participatory live coding. *PLoS Comput. Biol.* 16 (2020), e1008090. Issue 9. https://doi.org/10.1371/journal.pcbi.1008090

[5] Wirawan Purwanto, Bahador Dodge, Karina Arcaute, Masha Sosonkina, and Hongyi Wu. 2024. DeapSECURE Computational Training for Cybersecurity Students: Progress Toward Widespread Community Adoption. *The Journal of Computational Science Education* 15 (2024), 2–9. Issue 1.

[6] Wirawan Purwanto, Issakar Doude, Yuming He, Jewel Ossom, Qiao Zhang, Liwuan Zhu, Jacob Strother, Rosby Asiamah, Bahador Dodge, Orion Cohen, Masha Sosonkina, and Hongyi Wu. 2022. DeapSECURE Lesson Modules. https://deapsecure.gitlab.io/lessons/

[7] Wirawan Purwanto, Yuming He, Jewel Ossom, Qiao Zhang, Liuwan Zhu, Karina Arcaute, Masha Sosonkina, and Hongyi Wu. 2021. DeapSECURE Computational Training for Cybersecurity Students: Improvements, Mid-Stage Evaluation, and Lessons Learned. *The Journal of Computational Science Education* 12 (2021), 3–10. Issue 2.

[8] Wirawan Purwanto, Hongyi Wu, Masha Sosonkina, and Karina Arcaute. 2019. DeapSECURE: Empowering Students for Data- and Compute-Intensive Research in Cybersecurity through Training. In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning) (PEARC '19)*. ACM, New York, NY, USA, Article 81, 8 pages. https://doi.org/10.1145/3332186.3332247

[9] Wirawan Purwanto, Mohan Yang, Peng Jiang, Masha Sosonkina, Hongyi Wu, Jael Perales, and Kayla Curtis. 2024. T3-CIDERS Project Website. https://sites.wp.odu.edu/t3-ciders/

[10] Bo Zhu. 2015. A pure Python implementation of AES. https://github.com/bozhu/AES-Python.git