A Cyber Security Library – The need, the distinctions, and some open questions

Sidd Kaza, Department of Computer and Information Sciences, Towson University,
skaza@towson.edu

It is clear that in order to address the cybersecurity education and workforce crisis, the challenges are not just numerous but also inextricably linked. The least of which include a greater number of prepared faculty, effective curriculum, and infrastructure to host, use, and disseminate the curriculum. There is a demonstrated need for a cybersecurity digital library (DL) that will help address these challenges. The Cyber DL is similar to other curricular digital libraries in some respects (material quality, uptake, etc.) and unique in others (national security concerns, presence of damaging material – malware, material integrity issues, etc.). This idea paper articulates the need, the similarities, the distinctions and open questions, and provides some insights based on an ongoing Cyber DL project.

A Cybersecurity Digital Library – The need

Perhaps the greatest challenge to a successful digital library is the buy-in of the community behind it. For a cybersecurity digital library, this community includes academicians, industry, government standards and designation bodies, and the students who need the effective curriculum to contribute to our nation's workforce. Academia has taken advantage of the funding available from the National Science Foundation, National Security Agency, Department of Homeland Security, and other funding agencies available in the cybersecurity education arena. We have clearly reached a tipping point where there is effective curriculum to be had, only if there was a place to find it. There are early innovators responding to the need for curriculum sharing in cybersecurity education, such as CyberWatch, Department of Homeland Security (DHS), and SkillsCommons.org. There are similar efforts in computer science such as Ensemble, EngageCSEdu, NCWIT and in other STEM fields as well. The existing repositories offer several good features and a solid base on which to build, but there are several issues that need to be considered in the five-year horizon for a cybersecurity digital library to succeed.

A Cybersecurity Digital Library – learning from others

Vannevar Bush suggested the use of computers to retrieve information in 1945 (Bush 1945). The most recent surge in the term "digital library" came with the National Science Foundation funding research in the area through the Digital Library Initiatives through the nineties and into this century. There is a much cited formal framework focused on Streams, Structures, Spaces, Scenarios, and Societies to define digital libraries rigorously (Gonçalves et al. 2004) - Streams are sequences of items that describe static and dynamic library content. Structures are labeled directed graphs, that impose organization. Spaces are sets with set operations that obey certain constraints. Scenarios consist of sequences of events that modify states of a computation in order to accomplish a functional requirement. Societies are sets of entities and activities and the relationships among them.

A successful Cybersecurity Digital Library effort, has much to learn from the DL literature on what makes a "good digital library." There can be several quality indicators of the digital objects, metadata, collections, catalog, and services for a digital library. These include (Goncalves et al. 2007) accessibility, accuracy, completeness, composability, conformance, consistency, effectiveness, efficiency, extensibility, pertinence, preservability, relevance, reliability, reusability, significance, similarity, and timeliness. This is a rather long laundry list of quality indicators, and each is accompanied by metrics to measure them. As we build a Cyber DL, we will need to interpret and apply each of these to the new digital library.

A Cybersecurity Digital Library – Distinctions

There are several unique aspects and challenges to a Cyber DL that have not been explored in the digital library literature. In our work in building a prototype Cyber DL (www.clark.center) and working with the community, and beta-testers, we have identified the following issues (technical, policy, and social) that highlight the distinctions.

Complicated security policies – A Cyber DL will likely store cybersecurity curriculum that might provide the knowledge needed to cause malicious damage. One might argue, that such knowledge is found quite easily at other places on the web. However, this curriculum might be accompanied by pieces of Malware that will be used in sandboxed environments in the classroom (a rather common practice in security courses). Security policies need to be implemented to host, distribute, and sandbox this Malware. How do we ensure that an open Cyber DL does not become a “Dropbox” for Malware? How do we ensure that only qualified faculty have access to the materials?

Disclaimers and protection – Closely related with the previous policy issue, is the protection that a Cyber DL will need to have from potential damage the distributed content might cause. Does there need to be protection for the host – whether it be a university, a non-profit, or a private company?

Attacks from adversaries – As with any large-scale web application, security and availability would be a concern for the Cyber DL. However, producing cybersecurity professionals also contributes to our national security. Would a national Cyber DL become a soft target, needlessly attracting attention as it hosts curriculum that our CAE and other institutions use? If this indeed is an issue, what protocols and resources need to be in place to mitigate this risk and are they any different from other digital libraries?

Faculty incentives – Cybersecurity curriculum is challenging to build, deploy, and update. Though other disciplines might be similar, we can contend that cybersecurity learning materials will need to be updated more frequently and will require a dissemination plan so content consumers are not just notified but also involved in the maintenance of materials. If that is the case, the Cyber DL needs to include an incentive plan for content creators. Maybe a music subscription like plan (“the artist gets a small cut for each download”) or maybe a ‘tipping’ system (recommended at a recent workshop). In the age of Kickstarter, is a crowdsourced sustained funding source the way to go?

Storage, licensing, and dissemination – Several cybersecurity materials come with virtual machine (VM) environments that cater to the learning objects. Even with the seemingly endless storage capacity and bandwidth that we appear to have available, distributing VMs becomes a problem that scales very quickly. Cyber DL solutions will need to look at creative ways to not just store, but create a versioning for VM images, look at software licensing issues (and not become a "Dropbox" for pirated software), and look at bandwidth scaling very carefully so frivolous multiple downloads do not lead to escalating hosting costs. Should the Cyber DL consider partnering with a Cyber Range (Dark et al., n.d.) or maybe partner with a corporation (like Google) to donate storage and bandwidth?

The challenges in building a Cyber DL are many, but a discussion to answer some open questions will go a long way in making this digital library successful.

Acknowledgements

At Towson University, we are working on a pilot for a Cybersecurity Digital Library called CLARK (Cybersecurity Labs and Resources Knowledge-base, www.clark.center). CLARK is supported by the National Security Agency under NSA Grant H9830-17-1-0405. Though the language in this idea paper is the author's, some of the ideas are shared with Melissa Dark and Blair Taylor in their capacities with the NSA College of Cyber. To remove conflict of interest, neither Dark or Taylor reviewed this paper prior to submission.

References

Bush, V. 1945. “As We May Think.” *The Atlantic Monthly*, 1945.

Dark, M., S. Kaza, S. LaFountain, and Blair Taylor. n.d. “The Cyber Cube: A Multifaceted Approach for a Living Cybersecurity Curriculum Library.” In *The Colloquium for Information Systems Security Education (CISSE)*. New Orleans, LA.

Gonçalves, Marcos André, Edward A. Fox, Layne T. Watson, and Neill A. Kipp. 2004. “Streams, Structures, Spaces, Scenarios, Societies (5s).” *ACM Transactions on Information Systems* 22 (2). ACM: 270–312. https://doi.org/10.1145/984321.984325.

Goncalves, Marcos Andre, Barbara L Moreira, Edward A Fox, and Layne T Watson. 2007. “‘What Is a Good Digital Library?’ A Quality Model for Digital Libraries.” *Information Processing & Management* 43 (5): 1416–37. https://doi.org/10.1016/j.ipm.2006.11.010.

Author Bio

Dr. Sidd Kaza is the Chairperson and Associate Professor in the Computer and Information Sciences department at Towson University. He received his Ph.D. degree in Management Information Systems from the University of Arizona. His interests lie in cybersecurity education, data mining, and application development and he is a principal investigator on several cybersecurity education projects. He is also on the ACM Joint Task Force on Cybersecurity Education and is the recipient of the University System of Maryland Regent's Award for Excellent in Teaching. Dr. Kaza's work has been published in top-tier journals and has been funded by the National Science Foundation, National Security Agency, Department of Defense, Intel, and the Maryland Higher Education Commission.