

BitCurator.edu: Advancing Digital Forensics Education for Libraries and Archives

The University of North Carolina at Chapel Hill School of Information and Library Science (UNC SILS), in collaboration with the Educopia Institute, the BitCurator Consortium and the Council of State Archivists (CoSA), requests \$500,000 for a three-year Laura Bush 21st Century Librarian Program (LB21) research grant (7/1/2018 – 6/30/2021) to study and advance the adoption of digital forensics tools and methods in libraries and archives through professional education efforts.

Essential to advancing the **National Digital Platform** is the capacity of libraries and archives to acquire, manage and provide access to born-digital materials. There have been significant recent advances in applying digital forensics methods in libraries and archives to address this need. Measures include creating disk images to ensure the completeness, authenticity and availability of data; using cryptographic hashes for de-duplication and to identify, verify and authenticate materials; generating metadata to document the chain of custody; identifying and extracting contextual information; finding and locating sensitive information for further review, redaction or removal; building access environments that allow users to access data from disks using a web browser or emulation platform; and generating specialized reports to characterize contents of collections.

This work has been advanced by development and distribution of free and open-source software designed to support work in collecting institutions. The BitCurator project (2011-2014), funded the Andrew W. Mellon Foundation, developed the open-source BitCurator Environment, which is used actively by collecting institutions around the world to carry out the tasks listed above. BitCurator Access (2014-2016) and BitCurator NLP (2016-2018), also funded by the Andrew W. Mellon Foundation, have developed, documented and disseminated technologies that simplify and enhance access to born-digital data. The BitCurator Consortium (BCC) is an independent self-sustaining membership organization responsible for maintenance of the products of the three BitCurator projects, as well as advancing the community of professionals who use them.

Education is an essential component of advancing new professional practices, and there have been important advances in digital forensics education for library and archives professionals. UNC SILS has developed in-depth course offerings and has established a state-of-the-art digital forensics lab for use in research and teaching. Several other LIS programs (represented on the project Advisory Board discussed below) have also introduced hands-on digital forensics elements into their courses. On the continuing education front, there has been a growing set of classes on digital forensics offered through the Digital Archives Specialist (DAS) curriculum of the Society of American Archivists, as well as modules in CoSA's State Electronic Records Initiative (SERI) Institutes, and numerous specialized workshops targeting a variety of professional audiences. However, relatively limited education offerings continue to be a barrier to adoption of digital forensics in libraries and archives, despite advances in both tools and models for instruction. Relatively few LIS graduate education programs currently have the in-house capacity to thoroughly expose their students to these practices. The number of instructors providing continuing education is still quite limited; for several years, BitCurator.edu PI Lee taught the majority of them, though there has been some expansion of the instructor pool more recently.

The BitCurator.edu project will engage in research to investigate and support graduate-level LIS/preservation programs to implement educational offerings that make use of digital forensics tools and methods. This project will address two primary research questions:

RQ1: What are the primary institutional and technological factors that influence adoption of digital forensics tools and methods in LIS classes in different educational settings?

RQ2: What are the most viable mechanisms for sustaining collaboration among LIS programs on the adoption of digital forensics tools and methods?

The Advisory Board represents a diverse set of schools: Catholic University, Indiana University, New York University (Moving Image Archiving and Preservation Program), San Jose State University, University of Illinois, University of Maryland, University of Michigan, University of Texas, University of Wisconsin, Wayne State University. In addition to the project Advisory Board, we will also draw on a Professional Experts Panel (PEP) of individuals from the American Association for State and Local History (AASLH), CoSA, Digital Library Federation (DLF), LYRASIS, Massachusetts Institute of Technology (MIT), National Museum of American History, and the Southeastern Museums Conference (SEMC), who will provide input and guidance on all aspects of the project through periodic conference calls.

In addressing RQ1, the project team will work with members of the Advisory Board to test and implement several models for administering digital forensics education, ranging from use of dedicated classroom space, to relying on student's own laptops, to deployment of software functions over the Web. The project team will share guidance documents from these activities, which will serve as resources for education efforts elsewhere. RQ2 will be addressed through engagement with the BCC, Advisory Board and PEP on sustainability models, including (but not limited to) development of a new educational membership category for the BCC.

Work Plan: Year 1 will include data collection and needs assessment from advisory board members, and development of educational resources to test. Year 2 will include an in-person Advisory Board meeting and public symposium; collaboration with Advisory Board members on testing of educational resources; and development and deployment of two webinars targeting educators at other institutions. In years 2-3, in collaboration with CoSA, we will generate: 1) two webinars - one as an introduction to forensics for public records professions, and another on how to teach digital forensics tasks to others; and 2) a workshop (tentative venue: State Archives of NC) for records professionals to educate others about using digital forensics methods to advance electronic recordkeeping objectives. Year 3 will also include the concluding Advisory Board meeting (administered remotely), public symposium and generating of final project deliverables.

Dissemination: All products of the project team and will be made publicly available under a Creative Commons license, and promoted through various mechanisms, including social media, email lists, the project web site, publications and conference presentations. The BitCurator Consortium will host materials on its web site, and CoSA will do so through its Program for Electronic Records Training, Tools, and Standards (PERTTS Portal). The PEP will also help with outreach.

Outcomes and impact: This project will dramatically advance the **National Digital Platform** by 1) producing generalizable findings about factors influencing various educational implementation approaches and strategies for addressing them, 2) building the capacity of educational programs across the country to better prepare their students to effectively manage, preserve and provide access to born-digital materials, and 3) providing exemplars and guidance on experiential learning opportunities that can serve as models for other institutions.

Estimated Budget: The project budget requested is \$500,000. This includes a) a portion of PI (Christopher "Cal" Lee) effort, salary and benefits on this project; b) subcontract to engage Educopia which includes salary and benefits for Co-PI (Katherine Skinner), Project Manager and Community Lead (at the Educopia Institute), c) travel expenses for outreach and engagement; and d) travel and accommodations for one in-person Advisory Board meeting.