

A Graduate Certificate in Web Archiving

Michael L. Nelson, Michele C. Weigle, Sue C. Kimmel, Jessica Ritchie, Hongyi Wu

Summary: Old Dominion University (ODU) is the lead organization, and we seek \$98,361 (no cost share) for a one year planning grant to establish a graduate certificate in web archiving at ODU. The broad need we will address is the lack of a formal curriculum to support organizations that have web archiving as part of their mission. This project will have a broad impact by providing a multi-disciplinary training curriculum for both existing employees of libraries and archives, as well providing graduates for them to hire. The project category is “National Digital Infrastructures and Initiatives” and the project type is “Continuing Education”.

Statement of Broad Need: The web touches all parts of our lives, including education, commerce, government, and entertainment. There is a broad-based industry for advancing the core technologies necessary to construct web-scale information systems that power and enrich all areas of human interaction. Various computer science departments provide the research and training of graduates with the skills necessary to enter this industry, but these graduates often lack the library and information science background necessary for memory organizations. Harvard Library’s [Web Archiving Scan](#) (2016) recommended blending a traditional library focus on collection development with “smart technical development.” Our certificate represents such a synergy with computer science (Nelson, Weigle), librarianship (Kimmel), archives (Ritchie), and security (Wu) -- a synergy we believe is unique and will serve as a model for web archiving and general web science related collaboration elsewhere.

Noting a particular need for specialization in digital contexts, [BLS projects](#) “much faster than average” growth in the job outlook for archivists. However, finding graduates that have both the technical web skills and the archival and information science background is rare. The Internet Archive, Library of Congress, and other employers have to compete to hire unicorns or train their existing staff in one area or the other. The Web Science and Digital Libraries Research Group ([@WebSciDL](#)) in the Computer Science department at ODU is internationally recognized for its digital library and web archiving research as well as the quality of its graduates: in addition to conventional post-doc and tenure-track positions, our web archivists with PhD and MS degrees are finding employment at the Internet Archive, Library of Congress, Los Alamos National Laboratory, MITRE, Stanford, Berkeley, UCLA, and Emory, among others.

Despite our efforts, we cannot create enough MS and PhD graduates to fill the needs of these organizations. And for most existing staff at various libraries and archives, a complete graduate degree is also more than they need, or not a [realistic life option](#). If invited for a full proposal we will provide letters of interest in and support for our web archiving certificate from a mixture of universities, companies, museums, and libraries.

Project Design: The web touches all aspects of human discourse, and all the PIs are interested in the technology for its development and capture, theory of its organization and stewardship, and asserting its authenticity and integrity. With the unique faculty resources we have at ODU, we will innovate in both research and education in support of this vision. This project will harness the following synergies: 1) the internationally recognized [@WebSciDL](#) research group has grown to seven faculty, 2) the ODU Library & Information Studies program is the [first library science program in the state of Virginia](#) and is on the cusp of ALA certification, 3) ODU’s Center for Cybersecurity

Education and Research has just recently become the School of Cybersecurity. Within the scope of this planning grant, we will address the post-graduate training needs of the organizations listed above by building on the materials we co-created in the [CEDWARC](#) project and establishing *an on-line, four course (12 hour) graduate certificate in web archiving*, drawing from the faculty from Computer Science, Library & Information Studies, University Libraries, and the School of Cybersecurity. The planning grant will involve three phases:

Phase One (months 1--3): 1) All PIs will survey (lead by Weigle) existing USA graduate programs and certificate opportunities in Computer Science, iSchools, and Cybersecurity, and 2) perform (lead by Ritchie) a hiring needs survey of libraries and archives at the federal, state, and university level (e.g., skills they wish to hire (new staff) or grow (existing staff)).

Phase Two (months 4--12): Building upon Phase One: 1) develop a cross-listed capstone web archiving course (lead by Nelson, Kimmel, Wu): CS/CYSE/LIBS 650 - Web Archiving Theory and Practice (to be offered Fall 2022), 2) identify [possible electives](#) from the existing course offerings in Computer Science, Library and Information Studies, and Cybersecurity (Nelson, Kimmel, Weigle, Wu), and 3) propose possible new electives, including for a planned Archival Studies certificate (lead by Ritchie).

Phase Three (months 12+): Throughout phases one and two we will be self-evaluating and after the planning grant ends, all PIs will co-author a lessons learned document about both the certificate content and planning, as well as the meta-level interdisciplinary collaboration.

Diversity Plan: By offering this certificate online, we make it easier for remote and local working students to attend, including students that otherwise can not take time off from their jobs. By mixing topics from computer science and library science, we will increase gender diversity in both fields. While not an HBCU, ODU is a "[minority serving institution](#)" and has an enrollment that reflects the region. While this certificate offers graduate-level training without the full commitment to a MS degree, these courses would count toward a MS in the respective fields should students choose to continue.

Broad Impact: Web archiving straddles multiple disciplines, and the staff at libraries and archives rarely have the full range of skills necessary to perform these tasks. Organizations like the Internet Archive and Library of Congress have to internally train their staff, and lack a uniform curriculum to assist in this task. Based on our record of success and close ties with targeted organizations, we will provide the necessary training for federal, state, and university libraries and archives. In doing so, we will establish the first steps for web archiving as an emerging discipline unto itself.

Budget Summary: The budget will primarily support (\$64k, plus 32% overhead) course release and summer funding for all the PIs to support course development, multidisciplinary course strategy, and an employer needs survey. We assume travel will be possible by Fall 2021, so we will allocate \$10k to cover travel to the Library of Congress, NARA, Internet Archive, etc. We anticipate all PIs will be available for travel to DC (Norfolk is 3.5 hours away by car), but trips to California (e.g., IA) will involve fewer PIs.