Computational Thinking Integration into Archival Educators' Networked Instruction

Sarah A. Buchanan
iSchool
University of Missouri
Columbia, USA
buchanans@missouri.edu

Karen F. Gracy
School of Information
Kent State University
Kent, USA
kgracy@kent.edu

Joshua Kitchens
Archival Services and Digital Initiatives
Georgia Public Library Service
Atlanta, USA
joshua.kitchens@gmail.com

Richard Marciano
College of Information Studies
University of Maryland
College Park, USA
marciano@umd.edu

***Abstract*—This paper discusses the use of Computational Thinking (CT) in Archival Educators' instruction towards enhancing the training and professional development of the library and archival workforce to meet the needs of their communities, and enhancing digital collection management and access to information and resources through retrospective and born-digital content. Four educators share their teaching strategies aimed at modernizing the way digital LIS and computational education are conducted. Their goal is to create an active and engaged community of future archival practitioners, ready to tackle the digital records and archives future.**

Keywords—computational thinking, archival education, network of networks, digital archives, Jupyter Notebooks

I. Introduction

The vast majority of records that will be acquired by archives moving forward are being created computationally. Such born-digital acquisitions – and digitized analog archival holdings (an ongoing and expanding activity for most repositories) – require computational technologies for accessioning, managing, describing, accessing, compiling, mining and reusing digital content. Technologies for each task are advancing exponentially but there are still many unknowns as to their potential ethical, equity and social justice impacts. As federal agencies such as NARA are planning an even greater all-digital records future (Boyd 2021), there is a critical need to strengthen digital and computational literacy and training for future librarians, archivists and practitioners. Recognizing acute "skills and management gaps in libraries," IMLS funded the 2018 "Shifting to Data Savvy: The Future of Data Science in Libraries" (Burton et al. 2018) study which highlighted the need for greater automation in library work, the facilitation of computational research, and the need for library managers to understand the benefits of in-house data science skills.

IMLS also funded a number of initiatives to enhance the training and professional development of the library and archival workforce to meet the digital needs of their communities, including doctoral and master's-level training initiatives for the professional development of library and information science (LIS) workers. Such initiatives include Drexel's 2020-23 LEADING project[1] to create a nationwide cohort of LIS doctoral students and early to mid-career librarians, and our own 2020-22 Computational Thinking (CT) Piloting Network[2], a project aimed at prototyping a pilot to support MLIS training through a collaborative network of educators and practitioners that enables the sharing and dissemination of Lesson Plans and Computational Archival Science Educational System (CASES) Projects. IMLS has also supported public library staff to evaluate and improve CT programming for youth (ages 11-18) (2019-22 LG-14-19-0079-19). CT, described as a form of problem solving that uses modeling, decomposition, pattern recognition, abstraction, algorithm design, and scale (Wing 2006), is one of the foundations we build on.

In the last two years, fueled by COVID-19 distancing restrictions and advances in digital scholarship, cultural institutions such as GLAMs (Galleries, Libraries, Archives & Museums) have provided new forms of access[3] to the public through collections presented as data using the Jupyter Notebook platform. A short list of such datasets includes: Smithsonian Open Access, Austrian National Library, National Library of Scotland, Biblioteca Virtual Miguel de Cervantes, Bibliothèque Nationale de France, Library of Congress, British Library, Europeana, UGent Libraries, and Tim Sherratt's GLAM Workbench[4]. Growing 8-fold in number between 2018 and 2020, there are now nearly 10 million data-science notebooks[5] published, but few such notebooks are currently adapted for LIS audiences or the public. There is a real opportunity to train our students to become leaders in this GLAM space and help make available underutilized collections that will benefit the public.

[1] LEADING: LIS Education And Data Science Integrated Network Group, Jane Greenberg. See: https://cci.drexel.edu/mrc/leading/

[2] Piloting Network. See: https://ai-collaboratory.net/projects/piloting-network/

[3] Awesome Jupyter GLAM, Library Carpentry. See: https://github.com/LibraryCarpentry/awesome-jupyter-glam

[4] See: https://glam-workbench.net/

[5] As estimated by Alena Guzharina of the JetBrains Datalore team, see: https://blog.jetbrains.com/datalore/2020/12/17/we-downloaded-10-000-000-jupyter-notebooks-from-github-this-is-what-we-learned/

978-1-6654-8045-1/22/$31.00 ©2022 IEEE

II. Computational Thinking in Archival Science

The concept of Computational Thinking (CT) is being introduced to Science, Technology, Engineering and Mathematics (STEM) education. A definition of Computational Thinking for mathematics and science in the form of a taxonomy has been proposed that consists of four main categories: data practices, modeling and simulation practices, computational problem-solving practices, and systems thinking practices (Weintrop et al. 2016). In formulating that taxonomy, they draw on the existing Computational Thinking literature, interviews with mathematicians and scientists, and exemplary Computational Thinking instructional materials. Their work was part of an effort to infuse Computational Thinking into high school science and mathematics curricular materials. They argue for the approach of embedding Computational Thinking in mathematics and science contexts, present the taxonomy, and discuss how they envision the taxonomy being used to bring current educational efforts in line with the increasingly computational nature of modern science and mathematics.

That taxonomy consists of 22 Computational Thinking practices organized into four general categories of computational practices: (1) data practices, (2) modeling and simulation practices, (3) computational problem-solving practices, and (4) systems thinking practices (See Appendix A in Marciano, Jansen, & Underwood 2020). It has been shown that the 22 CT practices that have been identified as important in STEM education also were essential for performing archival practices when addressing digital records (Underwood et al. 2018; Marciano et al. 2019). It was also argued that such practices are essential elements of an archival science education in preparing students for a professional archival career (Underwood & Marciano 2019).

III. Archival Educators' Role

Archival educators are tasked with several responsibilities related to archival instruction – at the graduate level mostly but also for undergraduate students. They recognize the large role of practice in shaping the curricular components of archival courses, including students' desire for hands-on experiential learning activities like those in place at archival workplaces. Furthermore, educators regularly draw on such practitioner-based interactions with archives in introducing students to the types and kinds of actions expected to be performed on collections. Education, practice, and (practice-based) research therefore might be seen as interlocked and mutually-enhancing domains of archival science that benefit from direct contact with archival materials representing the breadth of the American experience. The cluster of collections assembled for the present work, named *"Re-presenting America,"* fulfills educators' imperative to educate prospective archivists on how to actualize the potential of collections as data (Thomas et al. 2019) to engage communities in novel reuse of their legacy materials, meeting new demands.

IV. Educator Network Contributions

The four members of an Educator Network (EN) worked over two years from Fall 2020 to Fall 2022 to achieve three goals: (1) develop computationally-enhanced lesson plans, (2) develop and conduct assessments of Computational Thinking (CT) knowledge, and (3) engage with Jupyter Notebooks, the related virtual MyBinder service, and (Python) programming learning resources. Simultaneously and iteratively, two members of an Educator Advisory Panel validated the goals and began to scope outputs to other programs. Though the goals remained consistent across the duration of the project, each one's direct application and appearance in classrooms varied by the specific focus and level of the course being taught. Project members agreed that lesson plan development should be driven by an educator's subject matter and course-content expertise. That resulted not in one lesson plan shared across the membership, but in multiple plans each tailored for use in one or more of the educator's courses. Ultimately, such lesson plans were launched and driven by the nature of practitioner collections engaged during the project and by the topics promised in respective course syllabi or weekly modules. Importantly, the practitioners themselves engaged with educators' courses in several ways including guest lectures and question-answering about past and future efforts with their archival collections. Practitioners' presence served to remind students of the people-driven and collaborative ethos characteristic of archival work environments (even when largely remote, as during the 2020-21 year). Educators also consulted a lesson plan template / strawman available open-access on the Computational Archival Science Educational System (CASES) website, https://cases.umd.edu/.

At the midpoint period, the EN in full participated in an Online Mini-Summit held on October 15, 2021 centered around the following research question: How can digitized archival collections and computational tools support P-20 educational goals? See: https://ai-collaboratory.net/2021/09/12/oct-15-2021-online-mini-summit/. EN contributions there revealed how a coordinated lesson plan effort from the Spring 2021 semester – making intentional use of three sequential datasets – resulted in both new storytelling about the individual and collective dataset(s), and new pursuits for practitioners managing that data that build on the EN students' findings, visualizations, and storytelling products. We detail educators' development and production of new instructional and research materials meant to be shared and reused by educators beyond the network. We tested CASES Projects over three initial semesters across 14 unique classes:

iSchools	Spring 2021	Fall 2021	Spring 2022	MLIS Topics
U. Missouri	• ISLT 9492: Data & Records Management	• ISLT 9490: Archival Practice	• ISLT 9491: Appraisal & Archival Systems	• Records Management • Appraisal
Kent State U.	• LIS 61095: Intro. to Digital Humanities	• LIS 61095: Archival Description	• LIS 61095: Intro. to Digital Humanities	• Digital Humanities • Description
Clayton State U.	• ARST 5110: Archives and the Web	• ARST 5300: Digital Preservation	• ARST 5100: Archives, Records and Technology • ARST 6950: Archives Capstone	• Preservation • Access
U. Maryland	• INST 742: Implementing Digital Curation • INST 443: Tools & Methods for Dig. Curation	• INST 448: Digital Curation Research in Cultural Big Data Collections • INST 341: Introduction to Digital Curation	• INST 747: Research in Advanced Digital Curation • INST 742: Implementing Digital Curation	• Digital Curation • Computational Thinking • Computational Archival Science

Fig. 1. 14 classes integrating CT taught by Educator Network members over initial 3 semesters of networked instruction.

A. U. Missouri: three graduate archival courses

The CASES website articulates five data practices that occur under the general banner of computational archiving practices (in addition to data practices, three other categories of practices are modeling and simulation, computational problem-solving, and systems thinking). The five data practices are collecting, creating, manipulating, analyzing, and visualizing data (DP1 – DP5 respectively). The key integration of CASES' scope of "project descriptions, lesson plans, and CASE files" demonstrating the teaching and learning of computational archiving at the university occurred during the graduate course ISLT 9492: Data & Records Management taught during Spring 2021, with resulting components later incorporated into two other archival courses in the subsequent semesters. Manipulating Data (DP3) and Visualizing Data (DP5) are the two that best match the above course's overall objective to develop *skills* for ensuring access and use of analog and digital formats. Of course, there exist other CT practices, thus, at both pre- (n = 27) and post- (n = 20, seven non-responders) Assignment intervals, students' knowledge of six such practices relating to archival work was assessed. The six CT skills assessed in that way are decomposition, pattern recognition, abstraction, algorithm design, modeling, and scale.

A Data Modeling assignment fulfills that objective through its emphasis on students' skills development manipulating and visualizing "real" data provided by an authentic agency. For that assignment, students worked in groups of three to closely examine a dataset, and corresponding data dictionary, for patterns embedded in the dataset as organized or digitally arranged. As part of that examination, students agreed upon unique questions to ask and then answer of their dataset by using any number of available tools for manipulation, e.g., OpenRefine, and visualization, e.g., RAWGraphs. Student groups produced highly original research questions and visualizations over the month-long duration of assignment work. Then, using MyBinder – to execute and interact with Jupyter Notebooks in the cloud without software installation – the educator converted the student groups' Computational Stories (assignments submitted) into open-access Jupyter Notebooks. Each Notebook makes use of the same dataset, named Index Card Data (as also in the below Section IV.D., though while there the focus was on personal experiences, the present work focused on group activity). The visualizations responded to diverse aims and interests: working timelines of a hunger strike provided a new tool of radical empathy for persons involved, a beeswarm plot sought out any relationships between types of "crimes" as documented in this and other datasets, and an undirected graph of two dozen individuals connected to a murder victim portrays the compromised nature of interpersonal connections in the Japanese-American internment camp Tule Lake.

The two MLIS topical emphases across the spring and subsequent courses are records management and appraisal. In the ISLT 9490: Archival Practice course of note, a student expanded their analysis from 138 index cards in two boxes (labeled as "lost and found") to 406 cards in 16 boxes. The student re-interpreted what they might have previously called "raw data" as instead multilayered material upon which archivists can create a story after the Stories and Stewardship Digital Curation Model (Kunda and Anderson-Wilk 2011), drawing original research conclusions having made new visualizations and language choices. In the iteration of ISLT 9491: Appraisal & Archival Systems, another student came to a similar view regarding reappraisal as had been the case during the earlier 9492 assignment: that space constraints are not always the primary concern or motivation, but instead, decipherment of an optimal arrangement based on any internal hierarchy of the collection is key to successful archival work when handling new digital archives acquisitions.

B. Kent State U.: two graduate library and archival courses

Students at Kent State University worked nonexclusively with two portions, named Entry Registry and Exit Registry, of the dataset shared during the collaboration by Denshō: The Japanese American Legacy Project based in Seattle, WA. Such data were digitized in 2015. MLIS students in the below courses gained knowledge in two topical emphases: digital humanities and (archival) description. Computational archiving theory and practice were incorporated into two graduate-level courses, one in digital humanities (offered in Spring 2021 and Spring 2022) and the other in archival description (taught in Fall 2021). Both courses were taught in asynchronous online mode.

The first course was LIS 61095: Introduction to Digital Humanities, taught in Spring 2021 and Spring 2022. Over the course of two semesters 22 MLIS students and one Ph.D. student were enrolled. The students had varied interests, including archives, digital libraries and projects, digital humanities, museum studies, and academic librarianship. Student knowledge of computational approaches ranged from complete newbie to basic preparation in scripting; few had more than rudimentary programming knowledge. Most students self-selected into the course; however, all had strong interests in learning how to use computational approaches to humanities research and teaching, and archives students were interested in taking a "deep dive" in archives data.

The course was structured into eight learning modules, ranging from one to three weeks in duration. Computational thinking was integrated into the course via the course structure, learning outcomes, course materials, learning assessments, and the term project. The final assignment served as a cumulative assessment of learning, as it required students to apply many of the foundational concepts to an actual dataset. As the assignment instructions stated, "This assignment is a culmination of concepts and skills that we will be learning throughout the class, applied to a real-life dataset. Over the course of the project, you'll familiarize yourself with the raw data, perform data clean-up and any manipulation required to create usable data suitable for analysis, conduct analysis on the data to identify patterns of potential significance, and create visual representations of those patterns. You will then create a digital project site that can display project outputs and provide context for your data." Students were given access to data from the Denshō Foundation, specifically the WWII-era Japanese internment camp (entry and) exit records that were part of Record Group 210 held by the U.S. National Archives (Records of the War Relocation Authority).

While students were given many opportunities for experiential learning during the course through the various technical exercises that accompanied each module (e.g., how to use optical character recognition (OCR) software, how to geocode data, how to do text analytics, how to create visualizations, etc.), the final project helped them "put it all together" and work semi-independently. While the instructor was always available as a guide, the students had the opportunity to design their path from raw data to the creation of a narrative that was underpinned with their analyses of the data. Throughout the final project, students were working with the data in a way that was informed by computational thinking.

The second course offered at Kent State was LIS 61095: Theory and Practice of Archival Description. 13 MLIS students were registered for the course in Fall 2021. Students were primarily interested in archival studies, with a few students interested in cataloging and museum studies who felt the description topic was applicable to their interests. As a 10-week course it was organized into eight modules; most modules were two weeks in duration.

The description course aimed to instruct students in 1) principles of provenance and original order; 2) hierarchical arrangement and description; 3) standards central to archival description including Describing Archives: A Content Standard (DACS), Machine Readable Cataloging (MARC), and Encoded Archival Description (EAD); 3) archival authority work, including Encoded Archival Context (EAC); and subject access for archival materials. Other related topics addressed include approaches to description of born-digital archival records, management of description programs, design of information systems for archival description, project management and cost analyses, and community-driven archival description such as tagging and reparative description.

Unlike the other Kent State course described above, in this course computational concepts were introduced primarily via the module on subject access for archival materials. Students explored the use of a semantic analysis engine to generate potential access points to collections from existing archival surrogates (i.e., finding aids). Future iterations of the course aim to incorporate computational thinking into other modules via activities such as tagging, transcribing, and optical character recognition, and via AI, ML, and additional Natural Language Processing approaches to the analysis of text found in archival records.

As introduced at the October 2021 Mini-Summit, Introduction to Digital Humanities as a new course was structured from the outset around the five data practices (listed in above Section IV.A), with three further surrounding modules focused on digital humanities (DH) work and institutions, CT as a DH meta-concept, and exhibits for showcasing DH projects. Working primarily with the (Entry and) Exit Registry datasets (with some other student-selected datasets allowed to supplement), students generated rich and interactive visualizations of displacement and relocation patterns otherwise not perceptible without computational analysis. Additionally, the Archival Description students connected the same data practices to research queries using a semantic analysis engine with online finding aids, altogether building technical skills through scaffolding, expert assistance, and hands-on practical experiences. Future iterations of each course are planned to incorporate use of Jupyter Notebooks to further cement understanding of computational thinking and provide an opportunity to build preliminary programming skills.

C. Clayton State U.: four graduate archival courses

The university is a Predominantly Black-serving Institution in the metro Atlanta area, offering a Master of Archival Studies (MAS) and a Minor in Information Studies which explicitly emphasizes digital archives and electronic records. Before this project, their home program Archives & Information Studies (AIS) centered curriculum on the Academy of Certified Archivists' "Role Delineation Statement" (ACA 2021) and Yakel and Torres' (2003) concepts of "archival intelligence." Such concepts define the core work of archivists and archival processes – including preservation, arrangement, description, and other aspects of archival work (Duff, Yakel, & Tibbo 2013). Those concepts were used to frame student assessment of learning and their level of understand or "archival intelligence." Mapping those concepts with computational thinking framework aided in aligning the two knowledge areas together. The present collaboration provided an opportunity to rethink globally how such formats were being taught at the graduate level and in an embedded context of social justice issues beyond the "archival silences" (Kitchens 2021) that are imposed through strictly technical approaches. Three such courses were reconceptualized, over the two-year timeperiod, by integrating CT into multiple key assignments each oriented toward archival intelligences (Yakel and Torres 2003) and appropriate selection of tools in practice. Given the dataset, named Camp Entry data, of center focus by the program during the collaboration, the two

MLIS topical emphases altogether characterizing the courses are preservation and access. Specific tools and practices are detailed below for ARST 5110: Archives and the Web, ARST 5300: Digital Preservation, and ARST 5100 Records and Technology (in the table, ARST 6950: Archives Capstone covers elements from each of the three).

- ARST 5300 Digital Preservation
  - The full-term 16-week course was taught in Spring 2020, 2021 and 2022 in the MAS Program at Clayton State University. Students were introduced to concepts for digital preservation and preservation management. Key practices (and tools) of the course are:
  - Preservation metadata (Premis)
  - Defining Technology Needs and Systems
  - Record Capture (Bagit, Data Accessioner, Bit Curator)
- ARST 5110 Archives and Digital Humanities, known as of 2022 as Archives and The Web
  - The full-term course was taught under the former title in Spring 2021 as a 16-week course, and under its present title in Summer 2022 as a nine-week summer course. Students were required to create digital humanities projects that explore or document marginalized groups. Students were allowed to choose any topic or software. Key practices (and tools):
  - Data Visualization and Analysis (Tableau, Story Maps)
  - Data Capture (Twarc)
- ARST 5100 Archives and Technology, known as of 2022 as Archives, Records and Technology
  - Taught in Fall 2020, Fall 2021, and Fall 2022, it is an introductory course in how digital systems and records interact. Students learn the fundamentals of digital records and various types of digital information. Key practices (and tools):
  - Introduction to Computational Thinking and Computational Archival Science
  - Jupyter Notebooks for data analysis
  - Virtualization and Emulations (VBox and Internet Archives)
  - Introduction to Programming (Python, PHP, Java)
  - Data Wrangling (OpenRefine)

D. U. Maryland: five undergraduate and graduate courses

The three MLIS topical emphases across five courses involved are digital curation, computational thinking, and computational archival science. Specific outputs, some based on the previously-mentioned Index Card Data and many made available for open-access learning and reuse, are detailed below for each course.

- INST 742: Implementing Digital Curation was taught twice:
  - **Spring 2021**: see https://ai-collaboratory.net/2021/05/06/may-6-2021_datathon/. The class concluded with a two-week datathon focused on a computational storytelling around the 1911 Charlotte, NC City Directory. The entire 14-week class was articulated around that single collection. Computational technologies included: graph databases, data visualization using Tableau, geospatial transformations using QGIS, cleaning and transforming using OpenRefine, and digitization management using ABBYFineReader.
  - **Spring 2022**: see https://ai-collaboratory.net/2022/05/09/may-5-2022-digital-curation-project-showcase/. Topics (and associated tools) explored included:
    - Archival Science concepts and workflows and Computational Thinking (CT)
    - Digitization management (ABBYYFineReader)
    - Cleaning & Transforming (OpenRefine)
    - Data Wrangling (Trifacta)
    - Clustering algorithms (Artificial Intelligence)
    - Text Processing through NLP and NER (GATE: General Architecture for Text Engineering)
    - Geospatial Transformations through: geocoding, geolocating, georeferencing, and vectorizing/tracing (QGIS, ArcGIS)
    - Data visualization (Tableau Storyline and Tableau Dashboard)
    - Network analysis through graph databases (Neo4j)
    - Digital Curation at scale
    - Virtual machines (UMD iSchool Virtual Computing Lab (VCL), Sandbox tools, Jupyter Notebooks)
- INST 443: Tools & Methods for Digital Curation.
  - An undergraduate class which started in the Spring of 2020 in person and in March became online due to COVID. Historical City Directories were examined through the lens of computational archival science.
- INST 448: Digital Curation Research in Cultural Big Data Collections.

- One of two undergraduate courses in the present scope taught in Fall 2021. A partnership was established with the Maryland State Archive's Legacy of Slavery Project, where the "Domestic Traffic Ads" collection was visualized and analyzed. That collection comes from newspaper advertisements placed by slave traders to buy or sell slaves throughout the state.

- INST 341: Introduction to Digital Curation.
  - **Fall 2021**: see https://ai-collaboratory.net/2021/12/14/dec-9-2021-digital-curation-showcase/. Practitioner Network members assisted in the development of topics showcasing social justice and computational storytelling using Jupyter Notebooks. 44 students from the UMD iSchool Undergraduate Information Science major showcased nine projects centered around four themes: 1. Redlining, 2. Japanese American Incarceration Camps, 3. Urban Renewal, and 4. Spelman College Alumnae Archives. Students in one of the 2nd theme projects went on to take part in a documentary first aired on April 9, 2022 by NHK World Japan, the international service of Japan's public media organization NHK. See: https://ai-collaboratory.net/2022/04/09/nhk-japanese-television-documentary/.

- INST 747: Research in Advanced Digital Curation.
  - **Spring 2022**: see https://ai-collaboratory.net/2022/06/18/june-18-2022-measuring-the-impact-of-urban-renewal/. Graduate students contributed to computational and policy work on the use of urban renewal data in Asheville, NC in order to provide evidence-based archival data and interfaces to inform the Asheville Community Reparations Commission in its 2022 effort to both evaluate the loss and define reparations with respect to Asheville's commitment to make amends for its destructive and discriminatory urban renewal program. A recent paper modeling the reparations work the course activities make possible was presented at DIGI-ARCHIVES-2022, the International Conference on Digital Archives, Big Data and Memory in Copenhagen, Denmark (Marciano et al., 2022).

V. Conclusions

The work illustrated across the above four educator contributions was part of a larger network referenced as the Piloting Network (consisting as well of practitioner, core, and advisory networks). See: https://ai-collaboratory.net/projects/piloting-network/team/. What has emerged from the hands-on experiential exercises mentioned is a validation of: (1) piloting computational networks, (2) applying computational thinking to library and archival science, (3) using Jupyter Notebooks for library and archival science education, (4) using machine learning (ML) for library and archival science automation, (5) using artificial intelligence (AI) in library and archival science teaching, and (6) developing infrastructure to support computational archival science (CAS). Through the educational network of networks, CAS is emerging as a new discipline (Marciano 2021). Not only has computational thinking been introduced into 'already-computational' archival coursework, as might have been our expectation at the start of the Fall 2020 semester. Instead, computational activities are no longer siloed into such culminating course(s) but successfully interrogated, referenced, and experienced in the traditionally introductory archives courses too. Such repeated engagements facilitate students' fluency with technologies established in archival workplaces and their willingness to seek out ongoing professional development as their education and career both continually advance into the future.

Acknowledgment

This research has been sponsored by the Institute for Museum and Library Science (IMLS) under two Laura Bush 21st Century Librarian (LB21) grants: a National Forum Grant [RE-73-18-0105-18] and a National Digital Infrastructures and Initiative (NDII) Project grant [RE-246334-OLS-20]. A recent IMLS grant extends this collaboration [RE-252287-OLS-22].

References

[1] Aaron Boyd. 2021. **NARA Plans to Expand Access to Digital Records Over the Next 5 Years**, Nextgov.com, Aug 5, 2021, https://www.nextgov.com/cxo-briefing/2021/08/nara-plans-expand-access-digital-records-over-next-5-years/184316/.

[2] Matt Burton, Liz Lyon, Chris Erdmann, and Bonnie Tijerina. 2018. **"Shifting to Data Savvy: The Future of Data Science In Libraries."** Project Report. University of Pittsburgh, Pittsburgh, PA, http://d-scholarship.pitt.edu/id/eprint/33891.

[3] Jeannette Wing. March 2006. **Computational Thinking**, Communications of the ACM, 49(3), pp 33–35, https://doi.org/10.1145/1118178.1118215.

[4] David Weintrop, Elham Beheshti, Michael Horn, Kai Orton, Kemi Jona, Laura Trouillel, and Uri Wilensky. 2016. **Defining Computational Thinking for Mathematics and Science Classrooms**, Journal of Science Education and Technology, 25(1), pp. 127–147, www.terpconnect.umd.edu/~weintrop/papers/WeintropEtAl_2015_DefiningCT.pdf.

[5] Richard Marciano, Gregory Jansen, and William Underwood. 2020. **"Developing a Framework to Enable Collaboration in Computational Archival Science Education"**, 2019 Society of American Archivists (SAA) Research Forum, Aug. 2, 2019, Austin, TX, https://www2.archivists.org/sites/all/files/Research_Forum%20_2019_Marciano_final.pdf.

[6] William Underwood, David Weintrop, Michael Kurtz, and Richard Marciano. 2018. **"Introducing Computational Thinking into Archival Science Education"**, 2018 IEEE International Conference on Big Data, 3rd CAS Workshop, Dec. 12, 2018, Seattle, WA, pp. 2761-2765. See: https://ai-collaboratory.net/wp-content/uploads/2020/03/1.Underwood.pdf.

[7] Richard Marciano, et al. 2019. **"Reframing Digital Curation Practices through a Computational Thinking Framework"**, 2019 IEEE International Conference on Big Data, 4th CAS Workshop, Dec. 11, 2019, Los Angeles, CA. See: https://ai-collaboratory.net/wp-content/uploads/2020/04/ReframingDC-UsingCT_final.pdf.

[8] William Underwood and Richard Marciano. 2019. **"Computational Thinking in Archival Science Research and Education"**, 2019 IEEE International Conference on Big Data, 4th CAS Workshop, Dec. 11, 2019,

Los Angeles, CA. See: https://ai-collaboratory.net/wp-content/uploads/2021/03/Underwood.pdf.

[9] Will R. Thomas, Benjamin Galewsky, Sandeep Puthanveetil Satheesan, Gregory Jansen, Richard Marciano, Shannon Bradley, Jong Lee, Luigi Marini, and Kenton McHenry. 2019. **"Petabytes in Practice: Working with Collections as Data at Scale,"** Data and Information Management, 3(1), pp. 18-25, https://doi.org/10.2478/dim-2019-0004.

[10] Sue Kunda and Mark Anderson-Wilk. 2011. **"Community Stories and Institutional Stewardship: Digital Curation's Dual Roles of Story Creation and Resource Preservation,"** portal: Libraries and the Academy 11(4), pp. 895-914. http://doi.org/10.1353/pla.2011.0047.

[11] ACA: Academy of Certified Archivists. 2021. **Role Delineation Statement for Professional Archivists**, The Academy of Certified Archivists. https://www.certifiedarchivists.org/role-delineation.

[12] Elizabeth Yakel and Deborah Torres. 2003. **"AI: Archival Intelligence and User Expertise,"** The American Archivist 66(1), pp. 51-78. https://doi.org/10.17723/aarc.66.1.q022h85pn51n5800.

[13] Wendy Duff, Elizabeth Yakel, and Helen Tibbo. **"Archival Reference Knowledge,"** The American Archivist 76(1), pp. 68–94. https://doi.org/10.17723/aarc.76.1.x9792xp27140285g.

[14] Joshua Kitchens. 2021. **"Engaging with Silences: Clayton State Master of Archival Studies Program's Approach to Teaching,"** International Journal of Information, Diversity, & Inclusion IJIDI 5(2), pp. 99-110, https://doi.org/10.33137/ijidi.v5i2.34818.

[15] Richard Marciano, et al. 2022. **"Promoting Archival Engagement through Computational Interventions"**, DIGI-ARCHIVES-2022: International Conference on Digital Archives, Data and Memory, final conference of the international network Digitization and the Future of Archives, Aug. 26, 2022, Copenhagen, Denmark. See: https://ai-collaboratory.net/wp-content/uploads/2022/08/Marciano-and-team_Paper_08-26-2022.pdf.

[16] Richard Marciano. Sep. 2021. **"AFTERWORD: Towards a New Discipline of Computational Archival Science (CAS)"**, in *Access and Artificial Intelligence: Working with Born-Digital and Digitised Archival Collections* (pp. 205-218), Lise Jaillant (Ed.), Bielefeld University Press. See: https://cup.columbia.edu/book/archives-access-and-artificial-intelligence/9783837655841.