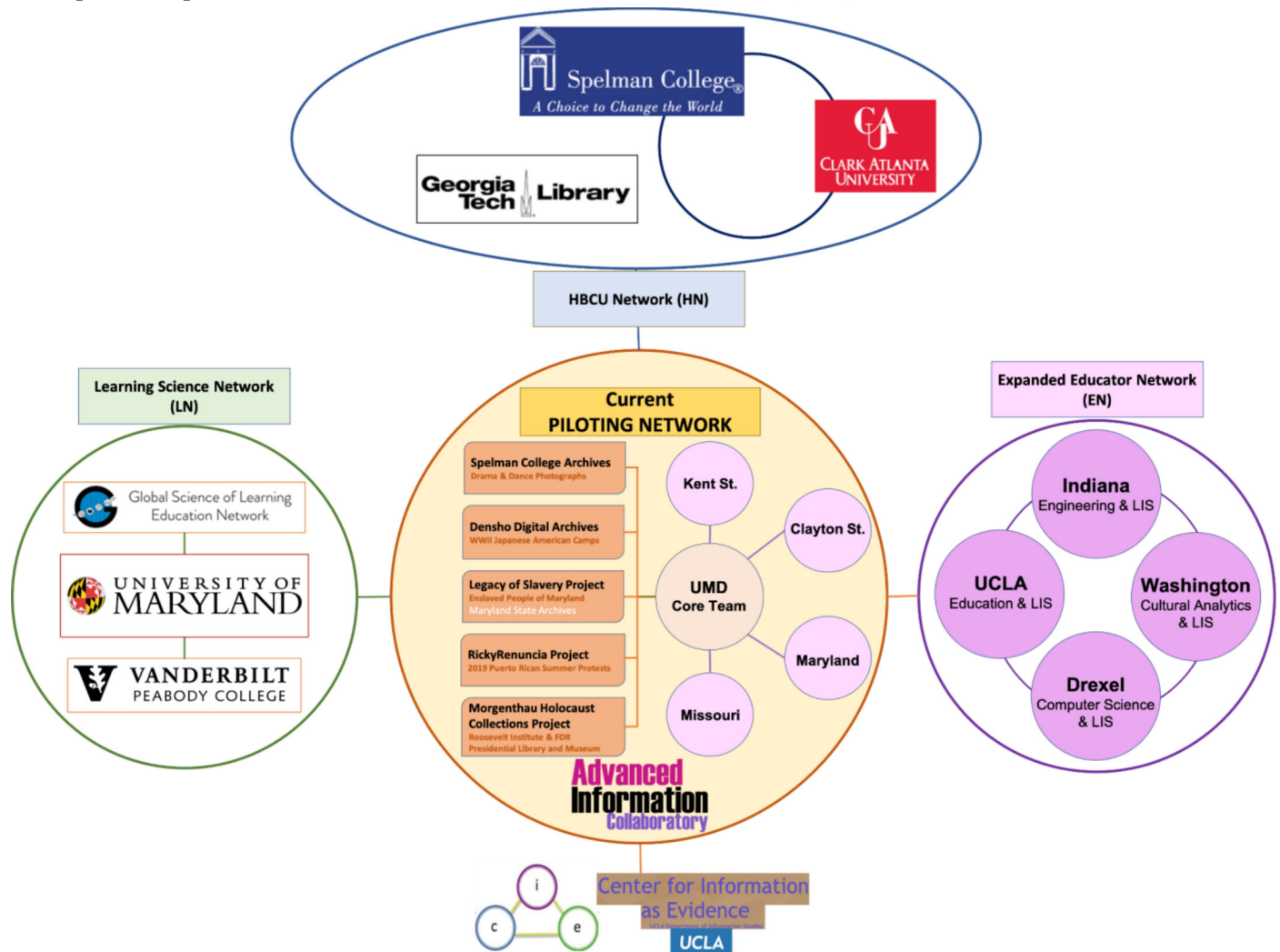


Launching the TALENT Network to Promote the Training of Archival & Library Educators w. iNnovative Technologies

The University of Maryland iSchool seeks a \$399,698 two-year Laura Bush 21st Century Librarian (LB21) Implementation Grant to scale up our existing “Piloting Network”¹ to a national network that will leverage the CASE Projects, Lesson Plans, Jupyter Notebooks, and Learning Education Platform we have developed and successfully tested in classroom settings, and help train the next generation of digital library and archives leaders. We plan to launch the TALENT Network, a collective group of 19 experts (including archivists, librarians, LIS educators, historians, learning scientists, cognitive scientists, computer scientists, and software engineers), focusing on four major objectives:

1. **Double our current Piloting Network** of iSchools through an **expanded Educator Network (EN)**. This will include multidisciplinary iSchools with a focus on “adjacent” fields (e.g., computer science, engineering, data science, and education), where there will be substantial opportunities for archivists and librarians with digital and computational skills.
2. **Engage HBCU students** (Historically Black Colleges and Universities) in the greater Atlanta community through an **HBCU Network (HN)** with a pilot program with two HBCUs: Spelman College and Clark Atlanta University (CAU), with coordination support from the Georgia Tech Library.
3. **Conduct curriculum development and review** through experts from a **Learning Sciences Network (LN)**.
4. **Address the social and ethical concerns** that arise from computational and algorithmic thinking through a partnership with the **UCLA Center for Information as Evidence (CIE)**.



¹ 2020-22 IMLS-funded “Piloting an Online National Collaborative Network for Integrating Computational Thinking into Library and Archival Education and Practice” – aka “Piloting Network”. See: <https://ai-collaboratory.net/projects/piloting-network/>

Over the next two years, the TALENT Network will increase and integrate the number of freely available **Lesson Plans**² (see *Supportingdoc1.pdf* for an example) and **CASE Projects**³ (see *Supportingdoc2.pdf* for an example) to equip MLIS educators and students to deal with the massive amount of information in electronic form. The idea behind Lesson Plans and CASE Projects is that they are powerful educational building blocks that can be adopted in a classroom setting at various levels – within a class, for use in a hands-on lab, across multiple classes, or even to construct an entire syllabus.⁴

More importantly, the TALENT Network will go beyond merely expanding the number of participating iSchools, but will also partner with leaders from the **Atlanta University Center Consortium (AUC Consortium)**, the oldest and largest contiguous consortium of African-American higher education institutions in the United States, in order to introduce HBCU students to the MLIS degree and LIS profession. Finally, the TALENT Network will make all archivists and librarians aware of their responsibilities to acquire and process information in a socially responsive, equitable, and ethical manner.

1. Project Justification

Our new TALENT (*Training of Archival & Library Educators w. iNnovative Technologies*) Network will address LB21 Program Goal 3: “*Enhance the training and professional development of the library and archival workforce to meet the needs of their communities,*” with primary objective 3.5: “*Enhancing digital collection management and access to information and resources through retrospective and born-digital content*”.

The vast majority of records that will be acquired by archives moving forward are being created in digital form. Digitization of existing analog archival holdings is an ongoing and expanding activity for most repositories. Computational technologies for managing, describing, accessing, compiling, mining and reusing digital content 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, 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](#)” 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 training initiatives for the professional development of library and archives professionals. These initiatives include Drexel’s 2020-23 [LEADING](#) project to create a nationwide cohort of LIS doctoral students and early to mid-career librarians, and our own 2020-22 [Piloting Network](#) at the U. Maryland, 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 CASE Projects. IMLS has also supported public library staff to evaluate and improve Computational Thinking (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 2 years, fueled by COVID-19 restrictions and advances in digital scholarship, cultural institutions such as GLAMs (Galleries, Libraries, Archives & Museums) have provided [new forms of access](#) to the public through collections presented as data using the Jupyter Notebook platform. A short list of these institutions 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](#). Growing 8-fold in number between 2018 and 2020, there are now nearly [10 million data-science notebooks](#) published, but few of these 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.

We propose a TALENT Network (*Training of Archival & Library Educators w. iNnovative Technologies*) to take our Piloting Network to the next level by scaling and accelerating its adoption and deployment. Beyond MLIS students, LIS educators, and archivists and librarians, our **target audiences** will also include: (a) educators in multidisciplinary iSchools with a connection to adjacent disciplines including engineering, computer science, cultural analytics, and education (in order

² For us, a **Lesson Plan** is a document instructors can use to structure the teaching of archival topics. It identifies the purpose of the lesson, associated Jupyter Notebook resources, archival topics addressed, relevant computational thinking topics, lesson objectives, pre/co-requisite knowledge, additional resources, typical plan and expected variations, and outcomes / assessment.

³ For us, a **CASE Project** is a Jupyter Notebook-based computational story that utilizes cultural data and is bundled as one or more interactive Jupyter Notebooks. [Jupyter](#) is “a free, open-source, interactive web tool known as a computational notebook, which researchers can use to combine software code, computational output, explanatory text and multimedia resources in a single document.”

⁴ A Lesson Plan can be constructed as a 20-minute, 45-minute, or even 90-minute class, and a CASE Project can be used for 3 to 4 class segments for example. A Lesson Plan can map to a single CASE Project or to many.

to attract students from these fields and expose our MLIS students to the opportunities in these spaces), and (b) undergraduate and graduate students in HBCU settings who might be attracted to the MLIS degree and LIS profession. This is a bold and broad approach, but elevating archival practices beyond their current uses demands it. Simply put, our goal is to modernize the way digital LIS digital and computational education is conducted by creating an active and engaged community.

2. Project Work Plan

2.1. Piloting Network Accomplishments

Our Piloting Network project has conducted extensive validation, testing, and dissemination. (results can be found on the Piloting Network website under: [presentations](#) and [publications](#)).

In our Piloting Network we emphasized developing Jupyter Notebooks around a unique cluster of collections we called the “*Re-presenting America*” collection, to emphasize its foremost significance and impact in exposing MLIS graduates to the full diversity of the American experience. The Piloting Network brought together four LIS programs (*U. Maryland, Kent State, U. Missouri, Clayton St.*) and five cultural institutions (*Spelman College Archives, Denso Digital Archives, Maryland State Archives Legacy of Slavery Project, Puerto Rican Spring Project, FDR Presidential Library Morgenthau Holocaust Collections Project*) to produce CASE Projects and an initial set of Lesson Plans that demonstrate how to support Master’s-level education in digital archiving. One of the unique features of the Piloting Network is the reliance on an archival science classification (using the Academy of Certified Archivists [RDS](#) or *Role Delineation Statement*) which lists seven major domains of archival practice and a series of tasks and related knowledge areas) and a computational thinking classification (using the [CT-LASER taxonomy](#) – *Computational Thinking to Library and Archival Science Education and Research*) which lists four major categories of CT concepts and a 22 CT Practices). These two classifications (called taxonomies) provide controlled vocabularies that allow the consistent mapping of archival activities to computational practices for any given CASE Project, hence the ability to teach in terms of shared concepts.

By the end of the Piloting Network project, we expect to produce 16 CASE Projects corresponding to 64 Jupyter Notebooks (remember that a CASE Project can be composed of 1 or multiple Jupyter Notebooks). At the time of submission of this proposal, the CASE Project breakdown was:

- I. 9 CASE Projects **completed** with 34 Jupyter Notebooks (linked to the CASES website) based on the collections of our cultural institution partners
- II. 5 CASE Projects **in progress** with 19 Jupyter Notebooks based on the collections of our cultural partners
- III. 2 Projects **in progress** with 11 Jupyter Notebooks based on general archival science topics

I. CASE PROJECTS (completed)	STATUS (as of 3/25/22)	COUNT
1. Legacy of Slavery Certificates of Freedom	Completed	3
2. Puerto Rican Spring	Completed	4
3. Japanese American WWII: Incident Cards & PII	Completed	7
4. Japanese American WWII: Biographical Framework	Completed	3
5. Japanese American WWII: Cards Solving Cases	Completed	7
6. Photomechanical Erasure in Archives	Completed	3
7. Redlining in Baltimore	Completed	3
8. Reading Punchcards	Completed	3
9. Web Scraping the Maryland State Archives	Completed	2
SUBTOTAL		34
II. CASE PROJECTS (in progress)	STATUS (as of 3/25/22)	COUNT
1. Historical City Directory for Charlotte NC 1911	In progress	5

2. Archival Storytelling: Redlining, Urban Renewal, Spelman Photo Archives, Japanese American WWII Camps	In progress	9
3. Legacy of Slavery Domestic Traffic Ads	In progress	1
4. FDR Presidential Library (using ML and AI)	In progress	3
5. WWI African American Photos (from NARA)	In progress	1
SUBTOTAL		19
III. CASE PROJECTS (based on archival topics)	STATUS (as of 3/25/22)	COUNT
1. <u>Digital Preservation</u> (Formats & Migration)	In progress (2/6 done)	6
2. Archival Description (APIs & JSON)	In progress	5
SUBTOTAL		11
TOTAL		64

At the time of submission of this proposal, the Lesson Plan breakdown was:

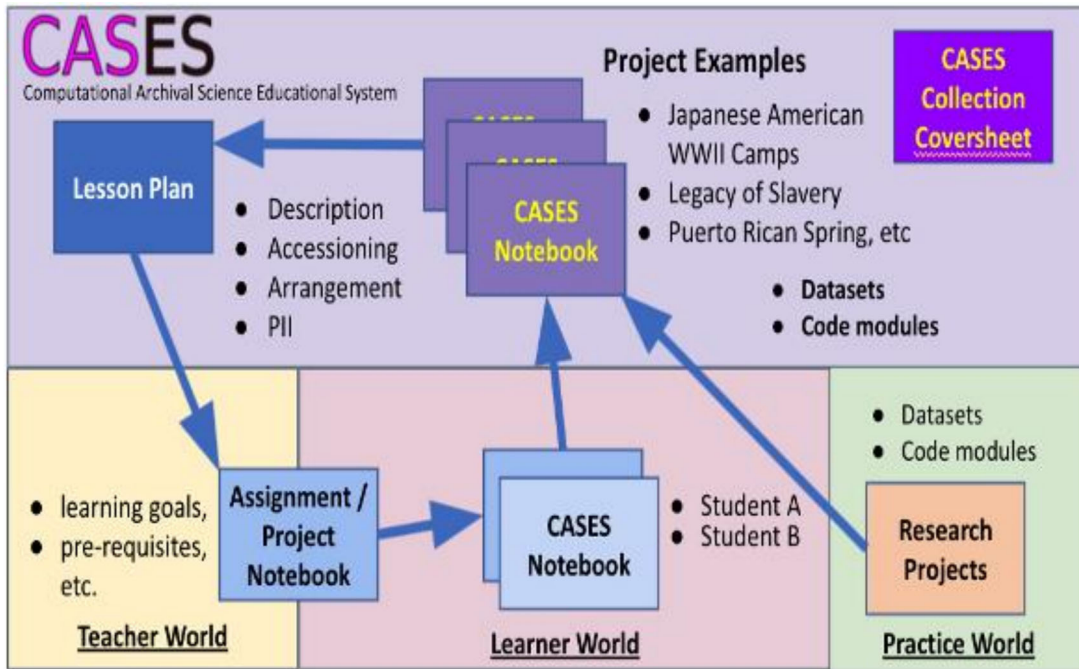
- 11 Lesson Plans with 6 archival topics:

LESSON PLANS	STATUS (as of 3/25/22)	COUNT
1. <u>Digital Preservation</u>	Completed	1
2. Arrangement / Description (<u>A</u>, <u>B</u>, <u>C</u>, <u>D</u>, <u>E</u>)	Completed	5
3. <u>Appraisal</u>	Completed	2
4. <u>Access and Outreach</u>	Completed	1
5. <u>Computational Archival Storytelling</u>	Completed	1
6. <u>Biographical Data</u>	Completed	1
TOTAL		11

We tested CASE Projects over 3 semesters across 15 classes through our current Educator Network of 4 iSchools:

iSchools	Spring 2021	Fall 2021	Spring 2022	MLIS Topics
Clayton State U.	• ARST 5110: Archives and the Web	• ARST 5300: Digital Preservation	• ARST 6950: Archives Capstone	• Preservation • Access
Kent State U.	• LIS 61095: Intro. to Digital Humanities	• LIS 61095: Archival Description	• LIS 61095: Intro. to Digital Humanities	• Digital Humanities • Description
U. Missouri	• ISLT 9492: Data & Records Management	• ISLT 9490: Archival Practice	• ISLT 9491 Appraisal & Archival Systems	• Records Management • Appraisal
U. Maryland	• INST 742: Implementing Digital Curation • INST 443: Tools & Methods for Dig. Cur.	• INST448: 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

What we learned from working with LIS students so far, is that they are capable of both “consuming” and “producing” Jupyter Notebooks. We envision the next iteration of the CASES Educational System in the TALENT Network to support the use, re-use, and creation of Notebooks by Teachers, Learners (undergraduate and graduate students), and Practitioners:



2.2. Project Objectives

We are inspired by two STEM-related initiatives that provide guidance to our project. The [CT-STEM](#) initiative at Northwestern University focuses on infusing computational thinking into high school science and mathematics curricular materials and provides a platform with free computational thinking lesson plan and assessments. The [Journal of Open Source Education \(JOSE\)](#) community focuses on developing software for assisting teaching and learning of open-source educational content and materials. The Editor-in-Chief of JOSE, Lorena A. Barba at GWU, has developed an inspiring [Teaching and Learning with Jupyter](#) handbook which discusses student and instructor benefits of using Jupyter Notebooks in terms of active learning and interactivity, and a catalog of pedagogical patterns that provides insights into suggested audiences, lecture formats, and learning goals. A CT-STEM community has emerged over the last seven years around these objectives.

In our project we wish to establish an iSchool community around CT-LASER, a mapping of Computational Thinking to Library and Archival Science Education and Research, where we help incubate and grow learning spaces that are uniquely centered around computational developments in library and archives collections, and their community of students, educators, and professionals. The TALENT Network thus proposes to build educator capacity, promote student engagement, learn from curricular experts, and address the ethical and social concerns that arise from algorithmic thinking:

1. **Expand to Multidisciplinary iSchools:** the current Educator Network (EN) comprises 4 LIS-centric schools (*U. Maryland, Kent State, U. Missouri, Clayton St.*). In this next phase, we will double the number of schools, adding partners from 4 multidisciplinary LIS schools that also emphasize computing, engineering, education, and cultural analytics. While these fields are adjacent to LIS, they are domains in which MLIS graduates are increasingly likely to work. These new nodes will include: Drexel College of **Computing & Informatics** (Prof. Adelaida Alban Medlock, Head of Undergraduate Affairs), Indiana U. Bloomington School of Informatics, Computing and **Engineering** (collaborator identified but to be confirmed), UCLA School of **Education & Information Studies** (Dr. Anne Gilliland), and U. Washington iSchool & **Cultural Analytics** (Prof. Melanie Walsh, author of [“Introduction to Cultural Analytics and Python”](#)).
2. **Engage HBCU Students in the Greater Atlanta Community:** we will launch a pilot program led by Spelman College with Clark Atlanta University (both HBCUs), with coordination support from Georgia Tech Library. This network will work with HBCU students in the greater Atlanta community to schedule a series of events (workshops, datathons), to develop additional Jupyter Notebooks and greater awareness of the MLIS degree and LIS profession. This will include an educational component that will assist in technical training through interactive discussions and introductory software skills development. The HBCU Network and partners include a team of notable professionals in fields including archives, computer and information science, and the humanities:
 - a. **Spelman College**, with Holly Smith, College Archivist, and Dr. Raquel Hill, Chair and Professor in the Computer and Information Sciences Department, and Assistant Professor to be named from C&IS.

- b. **Clark Atlanta University**, with Dr. Rico Chapman, Associate Dean, School of Arts and Sciences, Director, Humanities Ph.D. Program, Associate Professor of History.
 - c. **Georgia Tech**, with Dr. Aisha Johnson, Associate Dean for Academic Affairs and Outreach.
3. **Broaden Training through Alliances with Established Learning Sciences Networks:** we will develop broader alliances with established networks in *Learning Sciences*. The goal is to have our Lesson Plans and CASE Projects vetted and reviewed by learning science experts and build capacity by reaching out to additional collaborators in their own networks. Partners include:
- **Global Science of Learning for Education Network (GSoLEN)** (also funded through [NSF AccelNet-Design: Harnessing Global Science Networks to Accelerate Cultures of Learning](#)), and specializing in global development, learning, education, equity, and policy. The network brings together 600 members of GSOLEN and 100s of members within the Center for Applied Cognitive Science. The network is home to leading experts in learning and teaching and can provide input with respect to creating curricula to optimize student learning. The network is also aligned with large networks of teachers and professors around the world, allowing for further propagation of the work both directly and conceptually for the use of others. Prof. Andrea Chiba is a co-founder of GSOLEN and of the AccelNet team (UC San Diego, Cognitive Science and Program in Neuroscience). Her work focuses on learning and memory and notably many scientists and educators within the network focus on relevant topics.
 - **Vanderbilt Dept. of Teaching and Learning.** Prof. Rogers Hall, who is a curricular development expert. His work focuses on [bridging learning in urban extended spaces](#) by mapping archival media onto the city neighborhoods they describe and working with prospective high school teachers and developing learning sciences scholars.
4. **Address the Social and Ethical Concerns that Arise from Computational and Algorithmic Thinking:** we will partner with the [Center for Information as Evidence \(CIE\)](#) at UCLA, led by Dr. Anne Gilliland, in order to uncover the types of social and ethical issues that Computational and Algorithmic Thinking elicit in library and archival contexts.

Archives are increasingly critiqued as institutions whose holdings and practices favor high power records creators and users and disadvantage, marginalize and mischaracterize those who have historically not had such power. Given these facts, it is essential that archivists have the competencies to assess and respond to the opportunities, risks and ethical challenges posed by these technologies so that they may employ them in appropriate and liberatory ways in managing and providing equitable and enhanced access to their holdings. As yet, these competencies are not being systematically identified and developed in the archival field, even within graduate professional education.

2.3. Project Organization

We introduce the four networks our TALENT network is composed of a Core Network (CN), an Educator Network (EN), an HBCU Network (HN), and a Learning Science Network (LN):

- (1) The **HBCU Network (HN)** adds 5 new collaborators: 4 from 2 HBCUs (Spelman College Digital Archives and Computer & Information Sciences), and Clark Atlanta University; and 1 coordination collaborator from Georgia Tech Library.
- (2) The expanded **Educator Network (EN)** has 8 collaborators: 4 from our current network and 4 new collaborators from 4 multidisciplinary iSchools at: Drexel College of Computing & Informatics, Indiana U. Bloomington School of Informatics, Computing and Engineering, U. Washington iSchool & Cultural Analytics, and UCLA School of Education & Information Studies (which will double down as the lead on the social and ethical task force).
- (3) The **Learning Science Network (LN)** adds 3 collaborators from U. Maryland, the Global Science of Learning for Education Network ([GSoLEN](#)), and the Vanderbilt Department of Teaching and Learning & its [SLaMLab](#).
- (4) The **Core Network (CN)** at U. Maryland includes 5 experts from the iSchool in: Lesson Plan evaluation / learning science, digital archives, computational thinking, project management, and software engineering.

2.4. Project Activities

Development of Computationally-Enhanced Curricula (T1): Led by Richard Marciano

- **Task T1.1:** New multidisciplinary iSchool Educator Network partners (EN) and HBCU Network (HN) partners will benefit from the experience of the 4 original LIS-centric partners who will present existing Jupyter Notebooks for possible inclusion into their curricula.

- **Task T1.2:** Core Team (CT) will work with EN and HN to develop new CASE Projects and Lesson Plans for training in computing, engineering, education, cultural analytics domains, and for use with HBCU students and the public.
- **Task T1.3:** New Jupyter notebooks will include [LGBTQIA+](#) and [Chicano Studies](#) collections from UCLA.
- **Task T1.4:** Design a way to integrate Lesson Plans with CASE Projects. Right now, the CASES Learning Platform does not connect them.

Engagement with HBCU Students and the Greater Atlanta Community (T2): Led by Holly Smith

A pilot program led by Spelman College, Clark Atlanta University, with coordination support from Georgia Tech will explore working with HBCU students in the broader Atlanta community to schedule a series of events (workshops, datathons), providing exposure to Notebooks and MLIS-related professional opportunities:

- **Task T2.1:** Schedule and design a Virtual Workshop at the start of Year 1 (Fall 2022)
- **Task T2.2:** Explore virtual collaborations with HBCU students with the U. Maryland iSchool and Georgia Tech Library as partners (Spring and Fall 2023). The **Spelman Archives** would support the planning of workshops for students on data visualization, using computational data and would provide an overview of using archival collections for these activities, and also include a specific overview of the Spelman and AUC digital collections. They would be able to work directly with students, faculty, and other community stakeholders in facilitating in person research appointments to the Spelman Archives.
- **Task T2.3:** Plan a culminating research symposium (Spring of 2024). This will include the HBCU Network, partners, and community members. A discussion on the project and role academia plays in the Atlanta communities as a partner and collaborator will provide insight to additional ways for the contributors to seek influence for future projects. This will assist in garnering long-term success with collaborations and partners.

Participation in Learning Science Networks (T3): Led by Phil Piety

This network will use experts in Learning Science at Vanderbilt, the University of Maryland, and the University of California San Diego. Learning Science draws on Cognitive Science, Education Theory and Practice, and the use of technology mediated learning environments. This group will support curriculum development and review. They will make recommendations on how these products can be of value in other contexts.

- **Task T3.1:** A virtual presentation session will be internationally broadcast to the GSoLEN Network (organized by Andrea Chiba) and members with relevant expertise will be invited to form a working group. International partners may choose to engage in order to use this approach as a prototype for developing archival data and learning experiences surrounding their own countries' corrected histories. Here, data representation is a powerful tool for inclusion.
- **Task T3.2:** Vanderbilt University's Space, Learning & Mobility (SLaM Lab) engages faculty and graduate students in design studies of how physical and virtual spaces can better support learning. Members collaborate on research projects, design experimental teaching, and the interplay between physical and pedagogical environments. Rogers Hall will work with the TALENT Network to develop and study teaching activities with Lesson Plans and CASE Projects, integrated with teaching frameworks developed in SLaM work. These can be used and studied as teaching cases in courses offered at Vanderbilt to prospective high school teachers and developing learning sciences scholars.

Promotion of Ethical Information Access and Use (T4): Led by Anne Gilliland

Working with the UCLA Center for Information as Evidence (CIE), we will identify key social and ethical topics to be addressed through new CASE Projects and Lesson Plans for use and training purposes.

More than simply technical training, this project will educate archival students and practitioners to approach and deploy these tools with the intent of surfacing, evaluating and countering overt and latent biases and inequities in both the computational methods in the tools being used, and in the archival content and any associated metadata being processed. In approaching archival description and access, it will lay the basis for radical expansion not just of access, but also of analytical knowledge that will support a diversity of areas and concerns, including social and racial justice.

Attendant topics of biases and inequities resulting from computational treatments and data representation include:

- Risk exposure related to uncovering hidden narratives or exposing vulnerable or hidden individuals. Possibilities include the “right-not-to-be-exposed”.
- Computational storytelling that tells a story based on what is not explicitly in the records. Possibilities include the exploration of “computing and visualizing absences” and “making meaning out of silences”.
- Revealing deliberate tampering with digital content. Possibilities include the exploration of how such tampering can occur and be flagged.

- Surfacing bias that is “baked into” digital content or its metadata/representation. Possibilities include the exploration of how bias occurs and is manifested over time in these areas.
- Surfacing bias that is “baked into” algorithms and other computation processes. Exploration into how to hold algorithms, other ML and computational processes to account for inherent bias and progressive skew.

Tasks include:

- **Task T4.1:** to commence at the start of the project, will be a comprehensive review of all the existing Jupyter Notebooks and Lesson Plans to identify ethical information access and use issues that they raise or might potentially surface.
- **Task T4.2:** will be to identify new cases that either 1) address the same issues, but in different ways, or 2) address different or emerging issues (since this is a highly dynamic area). The latter will be based on input from a forum that will be convened of experts in algorithmic bias, digital ethics, and digital archives and libraries as well as a review of recent literature and cases.
- **Task T4.3:** will be to evaluate the extent to which and in what ways the new CASE Projects and Lesson Plans have indeed addressed the issues that were identified. This activity will also result in the development of a replicable evaluation framework that will be made available to developers of similar instructional materials.
- **Task T4.4:** will be to elicit risk principles through the development of a dynamically updatable framework of ethical and other risk principles and considerations that could be used in the library and archival fields for future computational developments.

Throughout this work there will be ongoing exploration of the issues raised in the first and second activities that will feed into the third and fourth activities.

Development of a Learning Education Platform (T5): Led by Greg Jansen.

The Learning Education Platform is a cloud-based environment that acts as a special kind of content management system and community space. Our Piloting Network project developed a prototype called the Computational Archival Sciences Education System (CASES). CASES allows stakeholders to visualize the various CASE Projects and Lesson Plans, download sample code and datasets, and upload contributions that add to the collection of resources for using these notebooks. CASES is based on a technology called Jupyter Hub, which allows the Jupyter Notebooks to be run in the cloud directly from our website (without having to download any execution software). In the TALENT Network, it will be expanded in several ways, including:

- **Task T5.1:** we will add Jupyter Notebook search functions to the CASES website. This will enable searching for keywords in either text or code sections across all notebooks, as well as searches for Notebooks by tagged topics from the Archival Science and Computational Thinking taxonomies.
- **Task T5.2:** the CASES platform will receive a complete software update to match the latest code and standards provided by the Jupyter Notebook project.
- **Task T5.3:** we will convert existing Lesson Plan documents to the Markdown text format and then publish these as CASES platform pages as well as make them searchable within the CASES platform. This will allow for the integration of Lesson Plans and CASE Projects.

Publication of a Jupyter Book with Computational Training Materials (T6): Led by Melanie Walsh

[Jupyter Book](#) is an open-source project for building beautiful, publication-quality books and documents from computational material, especially Jupyter Notebooks. We will hold a tutorial and series of sprint sessions to create a Jupyter Book from existing notebooks on the CASES platform.

Though Jupyter Notebooks are incredibly useful in their own right for exploring, teaching, and sharing code, they can also be challenging to work with in a number of ways. For example, a Jupyter Notebook file (.ipynb) cannot be opened and viewed on a computer that does not have the right software installed. While Jupyter Notebooks can be displayed on the web thanks to tools like *nbviewer*, which the CASES platform takes advantage of, the visual and aesthetic customization of these published notebooks is limited. Addressing these issues, Jupyter Book enables users to transform a collection of Jupyter Notebook files into readable, aesthetically pleasing, and fully customizable HTML files and to publish them as an easily navigable online book complete with a table of contents. The tool also enables the hiding or toggling of code, which makes the book’s pages more readable (especially for people who are unfamiliar with or intimidated by code), and it allows readers to download and share the book’s pages as PDF files. Jupyter Book thus enables Jupyter Notebook files to be shared with wider and more diverse audiences in a more polished and professional style.

- **Task T6.1:** will be a review and discussion of existing Jupyter Notebooks in the CASES platform, with the goal of deciding which notebooks might be best suited for inclusion in a cohesive Jupyter Book.

- **Task T6.2:** will be hosting a brief Jupyter Book tutorial and sprint session for starting a draft of the first Jupyter Book.
- **Task T6.3:** will be the transformation of remaining Jupyter Notebooks into the draft Jupyter Book and work on the aesthetic design and user interface of the Jupyter Book.
- **Task T6.4:** will be the dissemination of the draft Jupyter Book to team members and other relevant parties for feedback, suggestions, and discussion.
- **Task T6.5:** will be the revision and publication of a final Jupyter Book based on feedback from Task T6.4.

2.5. TALENT Network Expertise

HBCU Network (HN) & broader Atlanta community: See *Diversity Plan* section for more details.

- **HBCUs:**
 - **Spelman College:**
 1. **Holly Smith**, College Archivist, Spelman College.
 2. **Dr. Raquel Hill**, Chair and Professor, Computer and Information Sciences Dept., Spelman College. Dr. Hill’s primary research interests span the areas of trust, security, and data privacy. Her work on data privacy uses the semantics of data to create privacy preserving frameworks for text data.
 3. **Assistant Professor TBN**, Computer and Information Sciences Department, Spelman College.
 - **Clark Atlanta University (CAU):**
 4. **Dr. Rico Chapman**, Associate Dean, School of Arts and Sciences, Director, Humanities Ph.D. Program, Associate Professor of History, Clark Atlanta University. He is the Principal Investigator of a \$578K 2021-24 Andrew W. Mellon Foundation grant called “Building and Sustaining Digital Humanities at HBCUs.”
- **Georgia Institute of Technology Library:**
 5. **Dr. Aisha Johnson**, Associate Dean for Academic Affairs and Outreach, Georgia Tech. Formerly an assistant professor and program director in NCCU’s School of Library and Information Sciences and a lecturer at Clayton State University’s Master of Archival Studies program, Dr. Johnson also spent three years as Supervisory Archivist at the Jimmy Carter Presidential Library and Museum.

Educator Network (EN):

1. **Dr. Karen Gracy**, Professor Kent State U. School of Information. She has developed and delivered coursework in digital preservation and curation, moving image archiving, preservation and conservation of cultural heritage materials, and archival description and representation.
2. **Dr. Sarah Buchanan**, Assistant Professor, Library and Information Science (LIS), School of Information Science & Learning Technologies (SISLT), U. Missouri. She studies the creation and curation of data in the humanities, focusing in particular on data storytelling with audiovisual archives and provenance research methods. Being an archival educator, she is active in the Archival / Preservation Education ALISE SIG, in SAA CORDA and the Missouri Association for Museums & Archives, and in AERI and ASIS&T.
3. **Dr. Joshua Kitchens**, Director of the Clayton State Archival Studies program. He teaches a variety of courses at Clayton State that cover archival topics including Digital Preservation, Law and Ethics, Appraisal, and various special topics courses.
4. **Dr. Richard Marciano**, TALENT Network PI and Professor in the College of Information Studies at the University of Maryland. His research interests center on digital curation, sustainable archives, cyberinfrastructure, and big data. He is the 2017 recipient of the Emmett Leahy Award for innovation in records and information management and founder and director of the Advanced Information Collaboratory (AIC).
5. **Prof. Adelaida Alban Medlock**, Head of Undergraduate Affairs, Computer Science, Drexel University.
6. **Dr. Anne Gilliland**, Professor, UCLA School of Education & Information Studies, and Director of the Center for Information Studies (CIE).
7. **Dr. Melanie Walsh**, U. Washington iSchool & Cultural Analytics, [“Introduction to Cultural Analytics and Python”](#)).
8. **TBD**. At time of submission, we were in the process of finalizing the selection of a collaborator at Indiana U. Bloomington School of Informatics Computing and Engineering.

Learning Science Network (LN):

1. **Dr. Andrea Chiba**, UC San Diego, Cognitive Science Dept. [partnering with her Global Science of Learning for Education Network ([GSoLEN](#)) and specializing in global development, learning, education, equity, and policy]. Dr. Chiba’s work focuses on learning and memory and her center houses a network that is expert in

perceptual expertise and the value of teaching people to individuate within learning categories – an important factor in training the brain to eliminate cognitive bias.

2. **Dr. Rogers Hall**, Vanderbilt Dept. of Teaching and Learning [with a focus on [bridging learning in urban extended spaces](#) by mapping archival media onto the city neighborhoods they describe].
3. **Dr. Philip Piety**, Senior Lecturer and Learning Scientist at the University of Maryland iSchool. He is an expert in learning technologies and analytics and founder of the Maryland Education Digital Infrastructures and Analytics Lab (MEDIAL) that explores the tools that mediate pedagogy.

Core Network (CN):

1. **Dr. Richard Marciano**, Professor, UMD College of Information Studies, referenced earlier.
2. **Dr. Phil Piety**, Professor, UMD College of Information Studies, referenced earlier.
3. **Dr. Michael J. Kurtz**, project manager on our current Piloting Network grant and will continue in this role on the TALENT Network. Prior to this, he worked at the U.S. National Archives and Records Administration for 37 years as a professional archivist, manager, and senior executive, retiring as Assistant Archivist in 2011.
4. **Greg Jansen**, the AIC Collaboratory’s Principal Software Engineer and builds data repositories with new capabilities for computation and analysis. He has a focus on creating traceable data curation workflows that connect new data or holdings to evidence, through open workflows and chains of provenance.
5. **Mark Conrad**, digital archivist, educator, and researcher. He retired from the National Archives and Records Administration in October 2019 after 28 years. He worked with electronic records as an appraisal and accessioning archivist, program officer for e-records grants, educator on electronic records issues.

3. Diversity Plan

While we continue to focus on under-represented collections, our TALENT Network is concerned with diversity at multiple levels:

1. *Network*: All of the new partners in our expanded multidisciplinary iSchool Educator Network are either Hispanic, African-American, or Female and also represent new diverse disciplinary connections (computing, engineering, and cultural analytics, and education).
2. *Recruitment*: We will work closely with new partners in the HBCU Network, to ensure that we engage with HBCU students in the greater Atlanta area. The same will apply in the Educator Network.
3. *Collections*: In our Piloting Network we emphasized developing Jupyter Notebooks around a unique cluster of collections representing a wide range of the American experience, called “Re-presenting America.” We will continue to broaden these collections to include [LGBTQIA+](#) and [Chicano Studies](#) collections from UCLA.
4. *HBCU Network*: Collaborators include archivists, computer scientists, historians, and librarians who are deeply connected with their collections and have received accolades and national recognition for the quality of their scholarship and their leadership.

Through these partners, we will engage with the **Atlanta University Center Consortium (AUC Consortium)**, the oldest and largest contiguous consortium of African-American higher education institutions in the United States. The center consists of four [historically black colleges and universities](#) (HBCUs) in southwest Atlanta, Georgia: [Clark Atlanta University](#), [Spelman College](#), [Morehouse College](#), and the [Morehouse School of Medicine](#). The consortium structure allows for students to cross-register at the other institutions in order to attain a broader collegiate experience. They also share the [Robert W. Woodruff Library](#), a dual degree engineering program, and career planning and placement services.

4. Project Results

The TALENT Network aims at creating a durable, diverse, and multidisciplinary national community focused on developing archival and library educators and practitioners who can be future digital leaders. Recognizing acute “skills and management gaps in libraries”, we have established that there is an urgent need to change the MLIS culture with respect to computational and data science skills.

TALENT will enhance the training and professional development of the library and archival workforce to meet the needs of their communities while enhancing digital collection management and access to information and resources through legacy and born-digital content. Project results will include a rich library of repurposable and integrated Lesson Plans and CASE Projects based on an Archival Science and Computational Thinking framework, supported by an enhanced Learning Education Platform, a training manual, and a dynamically updatable framework of ethical and risk principles.

The deliverables will be developed and tested through an extended Educator Network of eight iSchools including four iSchools with adjacent domains of computing, engineering, education, and cultural analytics, an HBCU Network, and feedback from learning science experts in curriculum development.

Schedule of Completion

Aug. 1, 2022 – Jul. 31, 2024	NETWORKS				2022					2023					2024														
	CN	EN	LN	HN	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	A	
Task 1: Development of Computationally-Enhanced Curricula: Led by Richard Marciano																													
T1.1: Updates from old EN to TALENT																													
T1.2: Develop new HN & EN+ curricula																													
T1.3: Develop LGBTQIA+ & Chicano Studies curricula																													
T1.4: Design the integration of Lesson Plans & Case Studies																													
Task 2: Engagement with HBCU Students and the Greater Atlanta Community: Led by Holly Smith																													
T2.1: Joint Virtual Workshop																													
T2.2: HBCU student collaborations																													
T2.3: Research Symposium																													
Task 3: Participation in Learning Science Networks: Led by Phil Piety																													
T3.1: GSoLEN engagement (Andrea Chiba)																													
T3.3: Vanderbilt engagement (Rogers Hall)																													
Task 4: Promotion of Ethical Information Access and Use: Led by Anne Gilliland																													
T4.1: Comprehensive review																													
T4.2: Identify new bias-related cases																													
T4.3: Assessment & Framework																													
T4.4: Elicit risk principles																													
Task 5: Development of a Learning Education Platform: Led by Greg Jansen																													
T5.1: CASES search function																													
T5.2: CASES infrastructure update																													
T5.3: Implement curricula integration																													
Task 6: Publication of a Jupyter Book with Computational Training Materials: Led by Melanie Walsh																													
T6.1: Jupyter Notebook review																													
T6.2: Jupyter Book sprint session																													
T6.3: Design and prototype of Jupyter Book																													
T6.4: Dissemination of the draft Jupyter Book																													
T6.5: Publication of final Jupyter Book																													

Digital Product Plan

This project will create multiple digital products, including predominantly the Jupyter notebooks themselves, that combine code and narrative in a learning resource. It will also include lesson plans, surveys and other classroom assessment products, and website code that provides the CASES website. It may include minor ancillary scripts and code in support of learning taxonomies and the indexing of notebook content. We describe the plans for each product category in the sections below.

CASE Modules (Jupyter Notebooks)

Jupyter Notebooks, their accompanying lesson plans, and all supporting data, referred to as CASE modules, will be created by various team members. We expect to create between 30 and 100 such modules over the course of the project. These modules will be provided to the public as GitHub repositories and via the public-facing CASES website. This includes the facilitated delivery of notebooks and their related materials into the public-facing MyBinder.org JupyterLab notebook environment. The CASE modules will have copyright assigned to their creators or home organizations. The modules will be free for distribution and use under a Creative Commons license, requiring attribution only. CASE modules submitted to the core team for inclusion in the CASES website will undergo a review that includes appropriate copyright and license assignment as well as an assessment of cultural sensitivity and privacy issues. In cases where changes are needed, the core team will suggest some techniques for remediation of any identified issues. This may include suggested changes via Git, i.e. pull requests, which directly address the issue. (Remediation of this kind will necessarily also include the cleaning of the module's Git commit history.) Archival materials that receive attention and processing in Jupyter Notebooks will already be available to the public via the host institution. However, as a disruptive technology, the digital treatments and online access may raise new cultural sensitivity and privacy issues. These ethical concerns will be raised in all interactions with the Practitioner Network and continually examined as digital treatments are developed. Critical evaluation of ethical concerns is a core learning objective in this project's lesson plans. Therefore, we do hope to raise a number of such concerns in the lesson plans, explore the issues, and then demonstrate how they may be mitigated or avoided.

The goal of licensing and distribution of modules is to make them widely available for use in classrooms, to encourage communication among stakeholders, and to solicit the contribution of both feedback and enhancements. Free and active participation in enhancement activities is of central concern to the project, as we collaboratively develop learning resources and test them in classrooms. We think that as long as these modules are being taught in classrooms there will be active collaboration via CASES and GitHub to improve them and gather feedback.

CASES Website Code

The CASES website (<https://cases.umd.edu>) is a web application that is largely based on another web application, known as Notebook Viewer, developed by the Jupyter project. Our customizations have their copyright assigned to the University of Maryland at College Park and they are licensed for free distribution and use under the Apache 3.0 open source license, as required by the Jupyter Notebook Viewer code license for the upstream project. All of the code, with exceptions for secure configuration details, will continue to be made available via the public GitHub repository.

The CASES website code does not implicate any cultural sensitivity issues at this time. As mentioned in the previous section, CASES modules are reviewed prior to being included or featured in the website. The privacy concerns are minimal at this time. We do run an invitation-only JupyterLab on the cases.umd.edu site, for use by team members. However, no passwords are collected or stored there, as the authentication for that service is performed via GitHub's OAuth service. As this JupyterLab environment includes nascent CASE modules prior to review, it is accessible only to team members.

We aim to keep the CASES website updated and functional by depending whenever possible upon the upstream Jupyter Notebook Viewer project, which is more widely used and has contributors from around the world. This involves small, but periodic updates to the project based on changes in the upstream project's source code. These updates will be made by core team staff. The website service is also made more sustainable in the long term by use of the Docker environment for deployment. This makes the entire website stack more portable, if needed, to different infrastructure, such as cloud hosting environments.

Surveys and Classroom Assessments

The faculty using CASE modules in their classrooms may produce tabular classroom assessment data in spreadsheets or comma-separated values files (CSV). Such digital products involving students, in accordance with FERPA guidelines, will be cleansed of personally identifiable information prior to any publication. If any documents are found later to raise privacy concerns, we will immediately remove these from the public-facing project. Assessments will be shared internally with the team members of the various networks and may be published externally in aggregate or summary form, such as in journal articles.

Ancillary Code

We expect to create a small amount of ancillary scripts and configurations for back office use in data cleaning or the preparation content and taxonomies for publication. Where this code has lasting and public value, it will be published to GitHub with the Apache 3.0 license and with copyright assigned to the University of Maryland. We do not expect that this code will raise issues of cultural sensitivity or privacy, but we will conduct a review for these concerns prior to publication.

Organizational Profile

University of Maryland

Mission:

The University of Maryland College Park is a public research university, the flagship campus of the University System of Maryland, and the original 1856 land-grant institution in Maryland. The University of Maryland is dedicated to achieving excellence as the State's primary center of research and graduate education and the institution of choice for undergraduate students of exceptional ability and promise. With a commitment to diversity of faculty, students and staff, the University advances knowledge, provides outstanding and innovative instruction, and nourishes a climate of intellectual growth in a broad range of academic disciplines and interdisciplinary fields for the benefit of the economy and culture of the State, the region, the nation and beyond.

Service Area:

The Fall 2021 enrollment was 41,271 a total of graduate and undergraduate students. 42% of the population are minority students. 34.991% of students come from out-of-state and 65.009%% are Maryland residents. The University serves the state of Maryland as a premier research institution and reaches national distinction as ranking among the very best of public research universities in the United States.

College of Information Studies Maryland's iSchool in the Information Capital

Mission:

The College of Information Studies, Maryland's iSchool, engages in collaborative, interdisciplinary, and innovative research, teaching, and service. We educate information professionals and scholars, and we create knowledge, systems, and processes.

Service Area:

The iSchool offers Master's degrees in Library Science (MLS), Information Management (MIM), Human Computer Interaction (HCIM) and a doctorate degree in Information Studies. Per most recent admission data, 355 students are enrolled in the MLS program, 43 enrolled in the MIM program, 123 enrolled in HCIM program and 76 enrolled in the doctoral program. iSchool also offers Bachelor's degree in Information Science (BSIS) and the current number of enrollees is 1479. Approximately 42.39% of the total student body is female and 26.65% are underrepresented students. The iSchool has 45 Tenured and Tenured-track faculty, 35 Professional-track faculty, 54 staff and 113 adjunct faculty representing diverse subject areas in information studies. The iSchool serves the mid-Atlantic region.