Course Overview: This course focuses on mastering the complexity of research strategies and tools involved in the management and use of digital information in the Age of Big Data. The class will contain class lectures (including “pop-up lectures” as needed), class discussions, assigned readings, and extensive hands-on experience with student participation in digital curation projects. The research projects are focused around six major themes that will engage students in multiple arenas of research in Big Data. These are: community displacement, refugee narratives, movement of people, citizen internment, racial zoning, and cyberinfrastructure for digital curation. Project participants will have the opportunity to work with external stakeholders.

One example of a team project is:

- **Identifying relevant archival data** (digital and non-digital) from disparate record sources to analyze the impact of racially-based housing policies (i.e. redlining), and to document the nation-wide impact of these policies on neighborhoods across the country through the use of visualization and GIS technologies.

Another project example is:

- **Collaborating with the Maryland State Archives** in using digitized records involving Manumissions and Certificates of Freedom and Underground railway documentation to mine data and create finding aids to enhance user access.

A key element in the course is providing students with a strong understanding of agile management theory and techniques. This will assist students in developing the skills to manage projects in discrete segments and then be able to build on these segments in a coherent and technically feasible manner.
**Learning Outcomes:** Upon completing the course students will be able to:

- Demonstrate competency in managing the core digital curation lifecycle processes in the creation, use, preservation, and access of digital information.
- Develop and implement work plans to manage complex projects.
- Apply the technical, intellectual, and social awareness required to contribute to digital curation research projects.
- Construct effective research strategies to resolve or advance digital curation solutions to digital information challenges.
- Show insight and capability in applying digital curation tools in addressing research questions.
- Identify proper hardware and software to complete projects

**Course Prerequisites:** INST 604 would be desirable (Introduction to Archives and Digital Curation); permission of instructor can suffice.

**Course Requirements:** Students are expected to attend classes and actively participate in discussions sharing their own experiences and integrating information from lectures and assigned readings. Each student will participate in a major digital curation project and will demonstrate competence in contributing to team success and in addressing her/his own segment of the project.

**Grades:** Students will be evaluated on the basis of:

- **Participation in class discussions** 20%
- **Student contribution to team performance** 30%
- **Final Team product** 50%

A letter grade for each segment of student performance, and for the course, will be assigned with University and iSchool guidelines: A= Excellent [90-100]; B=Satisfactory [80-89]; C=Barely adequate [70-79]; D/F=Failure [<70].

**Team Project:** The research teams will meet throughout the semester and will present their products and conclusions in the final two class sessions. The deliverables due at the final sessions will be discussed in the first class session. Each team member will contribute to the team’s final product and will have an active role in the oral presentation in the final class sessions.
### Course Schedule:

<table>
<thead>
<tr>
<th>Session</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan. 30</td>
<td>Course introduction/Archival &amp; Digital Curation Lifecycle</td>
</tr>
<tr>
<td>2</td>
<td>Feb. 5</td>
<td>Overview of Potential Projects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Presentation (Bill Underwood): <em>Introducing Computational Thinking into Archival Science Education</em></td>
</tr>
<tr>
<td>3</td>
<td>Feb. 12</td>
<td>Selection of Semester Projects/Student Learning Objectives</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Presentation (Hamed Hanaee): <em>Jupyter Notebooks</em></td>
</tr>
<tr>
<td>4</td>
<td>Feb. 19</td>
<td>Team lab work</td>
</tr>
<tr>
<td>5</td>
<td>Feb. 26</td>
<td>Team lab work</td>
</tr>
<tr>
<td>6</td>
<td>Mar. 6</td>
<td>Team lab work</td>
</tr>
<tr>
<td>7</td>
<td>Mar. 13</td>
<td>Team project reviews (entire class)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>SPRING BREAK – MARCH 16 – MARCH 24</strong></td>
</tr>
<tr>
<td>8</td>
<td>Mar. 24</td>
<td>Team lab work</td>
</tr>
<tr>
<td>9</td>
<td>Apr. 2</td>
<td>Team lab work</td>
</tr>
<tr>
<td>10</td>
<td>Apr. 9</td>
<td>Team lab work</td>
</tr>
<tr>
<td>11</td>
<td>Apr. 16</td>
<td>External stakeholder feedback (entire class)</td>
</tr>
<tr>
<td>12</td>
<td>Apr. 23</td>
<td>Team lab work</td>
</tr>
<tr>
<td>13</td>
<td>Apr. 30</td>
<td>Final project reports (entire class)</td>
</tr>
<tr>
<td>14</td>
<td>May 7</td>
<td>Final project reports (entire class)</td>
</tr>
<tr>
<td>15</td>
<td>May 14</td>
<td>Course Recap</td>
</tr>
</tbody>
</table>

**Course Materials:** Course administrative documents, class or project assignments, and assigned readings will be accessible in the class ELMS site at the time the course occurs.
Readings:

2. Data Curation Network: https://sites.google.com/site/datacurationnetwork/
5. The “Computational Archival Science (CAS) Portal”. Link: http://dcicblog.umd.edu/cas/
11. TNA and KCL Workshop on Computational Archival Science: Automating the Archive, Sep. 7, 2018, The National Archives, Kew, UK. See: http://dcicblog.umd.edu/cas/9-7-2018_uk-nara-kcl_cas-workshop/
   - *The Jupyter Notebook is an interactive computing environment that enables users to author notebook documents that include:* - Live code - Interactive widgets - Plots - Narrative text - Equations - Images – Video.
   - *These documents provide a complete and self-contained record of a computation that can be converted tovarious formats and shared with others using email, Dropbox, version control systems (like git/GitHub) or nbviewer.jupyter.org.*

17. Case Studies on Teaching with Primary Sources (TWPS). See; https://www2.archivists.org/publications/epubs/Case-Studies-Teaching-With-Primary-Sources


From: https://dcicblog.umd.edu/cas/ieee-big-data-2018-3rd-cas-workshop/

• Paper #1 Introducing Computational Thinking into Archival Science Education
• Paper #2 Automating the Detection of Personally Identifiable Information (PII) in Japanese-American WWII Incarceration Camp Records
• Paper #12 A Case Study in Creating Transparency in Using Cultural Big Data: The Legacy of Slavery Project
• Paper #13 Jupyter Notebooks for Generous Archive Interfaces


Readings will be customized based on the particular project the students are working on. For example, a project involving data scrubbing, loading a graph database, formulation of network queries, will have related references on research methodologies and tools, whereas a project involving crowdsourcing, R-programming, Python and database building will call for a separate set of references, etc.

Attendance: Attendance in class is required. If you are unable to attend class for reasons of illness, religious observance, participation in University activities at the request of University authorities, or compelling absences beyond the student’s control, please let the instructor know [prior to class for non-emergencies].

Classroom Environment: The classroom environment should be professional and respectful. Punctual arrival at classes contributes to the smooth operation of the class and the quality of the learning experience. Late arrivals, early departures, and wandering in and out of the classroom are disruptive and distracting. If you know you will be unavoidably detained or must leave early, please let the instructor know in advance. Please turn off or mute all phones and other communication devices during each class session. If you use your laptop in the classroom, limit the usage of the computer to course-related reasons [i.e., taking notes].
**Academic Integrity:** The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Society. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student, you are responsible for upholding these standards for this course. The Code of Academic Integrity prohibits students from cheating on exams, plagiarizing, submitting fraudulent documents, forging signatures, submitting the same paper for credit in two courses without authorization, and buying papers. It is very important for you to be aware of the consequences of academic dishonesty. Instances of any suspected academic dishonesty will be reported and handled according to University policy and procedures. For more information on the Code, visit [http://www.shc.umd.edu](http://www.shc.umd.edu)

Plagiarism is of particular concern in the networked digital environment. Students must write their essays and assignments in their own words. Whenever students take an idea or a passage of text from another author, they must acknowledge their source both by using quotation marks where appropriate and by proper referencing using footnotes or in-text citations. For further information about proper citation of sources, consult the UMD Libraries website at [http://www.lib.umd.edu/guides/honesty.html](http://www.lib.umd.edu/guides/honesty.html) and [http://www.lib.umd.edu/PUBSERV/citations/index.html](http://www.lib.umd.edu/PUBSERV/citations/index.html).

**Extensions:** Late submissions of written assignments [including the final exam] will carry a penalty unless prior arrangements are made with the instructor. If an extension is granted, the work must be submitted within the extension period to avoid grade penalties. Unexcused delays in submission of the assignment will result in a reduction of the grade by one category for each day the paper is late; for example, a paper that would have received a B+ if submitted on time will receive a B if it is submitted a day late, a B- if it is two days late, and so on.

**Students with Disabilities:** Students with disabilities who require academic accommodations must inform the instructor of their needs and provide written documentation about the appropriate academic accommodations from Disability Support Services [http://www.counseling.umd.edu/DSS](http://www.counseling.umd.edu/DSS) at the beginning of the term.