

# IST676 Digital Curation

## Spring 2022

Instructor	Jian Qin	Phone	(315)443-5642
Office	226 Hinds	E-mail	jqin@syr.edu
Office Hours	By appointment	Time & Place:	120 Hinds Hall

---

### Description:

Introduces key concepts and techniques in digital curation across humanities, social sciences, and sciences and issues in the technical, service, and social dimensions of curating digital assets.

### Additional Course Description:

This course introduces key concepts revolving around digital data and their curation and user services, with a focus on the technical, service, and social/political/legal dimensions. The technical dimension covers the design and development of digital repositories that store, organize, and manage digital data. The service dimension discusses the infrastructure and institutional support for the use of digital repositories for learning, research, leisure, and other purposes. The social/political/legal dimension includes the external forces that will shape the production and use aspects of digital curation and data services. The course will examine topics such as digital data/assets and repositories, principles, tools, and policies for digital data services, data management plans, and the political and legal forces that may shape the digital curation environment. Students will incorporate these lessons into a term-long project – the planning and design of their own digital curation project.

Students will gain a thorough grounding for understanding, evaluating, and working with a wide variety of digital archives/repositories by combining individual and/or group assignments that will give students opportunities to investigate some of the specific issues of their interest. Lectures, hands-on exercises, discussions, short papers and a term-long project will provide experiences and learning opportunities for pursuing topics in greater depth and gaining proficiency in communicating technical and pragmatic issues surrounding digital curation and services.

### Audience:

Master graduate students and professionals interested in working in areas of digital curation, digital assets management, digital repositories, cultural heritage, digital preservation, archives, and research data management.

### Credits: 3

### Bibliography/ Texts / Supplies – Required:

Banerjee, K. & Reese Jr., T. (2019). *Building Digital Libraries: A How-To-Do-It Manual for Librarians*. Chicago: ALA Neal-Schuman. (eBook available in SU Library)

Bryant, R., Fransen, J., de Castro, P., Helmstutler, B., & Scherer, D. (2021). *Research Information Management in the United States: Part I—Findings and Recommendations*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/8hgy-s428>.

Cameron, F. (2021). *The Future of Digital Data, Heritage and Curation in a More-Than Human World*. Routledge. (eBook available in SU Library)

## IST676 Digital Curation Spring 2022

Harvey, D. R. & Oliver, G. (2016). *Digital Curation* (Second ed.). ALA Neal-Schuman, an imprint of the American Library Association. (eBook available in SU Library)

Lawrence, G.W., Kehoe, W.R., Rieger, O.Y., Walters, W.H., & Kenney, A.R. (2000). *Risk Management of Digital Information: A File Format Investigation*. Washington, D.C.: Council on Library and Information Resources. <https://www.clir.org/wp-content/uploads/sites/6/pub93.pdf>

Additional readings will be provided in the weekly schedule.

### Expectations and Course Specific Policies

**Participation:** Attending class and participation in class activities is expected, exactly as it would be on the job. If an emergency or illness occurs, have someone notify the course instructor as soon as possible—even if you are out of town. Absence from class for two weeks without notice will automatically result in a 20% deduction in your participation grade. Too many absences are sufficient to cause the lowering of the final course grade. Exceptions will be made for emergencies and other extenuating circumstances, provided they are verified by appropriate documentation that is received no later than one week after the absence(s).

**Conduct of discussion:** It is expected that students will behave professionally both in language and attitude when responding to discussions. Public disparagement of your fellow students in this course is unacceptable and may result in disciplinary action. Additionally, discussions should model your ability to think critically about course topics and articulate ideas clearly. Responses should be explanatory. If you have any concerns regarding the suitability of a post, please contact the instructor.

**Communication:** Detailed effort is put in to ensure every class is as meaningful as possible and worth attending. Students will be responsible for all materials covered, handed out (virtually), announced, and so on in class unless told otherwise. Attempts will be made, however, to place important announcements in Blackboard.

**Assignments and grading:** Every attempt will be made to return assignments in a timely fashion. Assignments are due at the time indicated in the weekly schedule, unless specified otherwise, and will be annotated with grading rubrics. Late work will be accepted for two days only after the due date, with a 5% penalty per day. This is to facilitate the timely return of graded assignments with answers.

Each assigned work will be graded on the scale as specified for the component (e.g., each exercise will receive a maximum number of 15 points), which will be summed at the end of the semester. Assignments meeting the basic requirements will be considered as a B grade work. To earn an A grade, the assignment must show evidence of extra effort.

An Incomplete grade, **I**, can be given only if the circumstances preventing the on-time completion of all course requirements were clearly unforeseeable and uncontrollable. If an Incomplete is required, a written contract must be completed that specifies the nature of the missing work, the date it will be completed, and the default grade that will be given if that deadline is missed.

It is unethical to allow some students additional opportunities, such as extra-credit assignments, without allowing the same options to all students.

Failure to complete any course requirement will result in a course grade of **C** or lower, regardless of the grades received in other components.

## IST676 Digital Curation Spring 2022

To discuss a grade, arrange for a private meeting (via phone or video conferencing) in which you identify the sources of your concern. It is important to provide at that meeting the relevant materials (e.g., marked papers). Except for extraordinary circumstances, no appeal for an individual assignment or project will be considered later than two weeks after the graded assignment was returned. For final grades, no appeal will be considered after two weeks of final project submission date.

This syllabus (including course requirements, due dates, etc.) may be changed with sufficient notice via SU email. If you have any type of disability that may require additional time or special consideration, please let me know *at the beginning of the course*.

### Learning Outcomes

Upon the completion of this course, students will be able to:

Course learning objectives	Program learning outcomes
1. Understand digital curation as well as their emerging roles in our learning and research environment	Connect, engage, and collaborate with users and their communities through the provision of inclusive resources, services, and spaces.  Demonstrate information literacy and technological agility.
2. Gain foundational ability to apply theories and technologies in the management of digital data and repositories	
3. Analyze and evaluate how digital curation services operate and meet user needs/requirements	
4. Practice and promote principles in data ethics and equitable and just access to digital data and services	Advance information equity and justice by applying professional ethics, values, and standards to work towards a just and equitable information society.
5. Learn how to plan a digital curation project and work productively and creatively	Design and innovate to create equitable, just, and engaging information artifacts, including services, systems, spaces, resources, and technologies.

### Assessment

Your course performance is assessed based on your ability to develop, utilize, and articulate your ideas concerning the topics covered in the course.

Coursework for IST676	Course Learning Objectives
<b>Labs</b> (5 x 5 points = 20 points, 25%) 1. Data documentation 2. Collections as data 3. Design and creation of curation-friendly data 4. Creating open datasets 5. OpenRefine exercise	1, 2, 4
<b>Case studies</b> (non-graded, class participation) 1. Decision on data repository solutions 2. Digital preservation 3. Linking data for discovery	3, 4
<b>Assignments</b> 1. Planning a digital data project (10 points, 10%) 2. The biggest problem (10 points, 10%) 3. Project report (key assignment, 40 points, 40%)	5

## IST676 Digital Curation Spring 2022

Class participation (15 points total: 5 points for attendance, 10 for participation, 15%) will be evaluated by

- a. Evidence of having read assigned readings.
- b. Asking meaningful and relevant questions.
- c. Volunteering answers and ideas.
- d. Active involvement in group activities and in-class exercises.
- e. Attendance is recorded and is included as a part of the participation grade.

### Grading Table\*

Grades	Grade Points/ Credit	Total Points	Notes
<b>A</b>	<b>4.000</b>	95–100	<b>A</b> = <i>Superior scholarship</i> through critical thinking, exemplary products, positive and supportive interactions with colleagues, and sustained active participation across course activities. <b>B</b> = <i>Average performance</i> on all assignments; this graduate standard indicates that the work was <i>well done</i> , complete, met stated criteria, represents a strong professional effort, and was turned in on time. <b>C</b> = <i>Below average</i> performance. <b>F</b> = <i>Failure</i> .
<b>A–</b>	<b>3.667</b>	90–94.9	
<b>B+</b>	<b>3.333</b>	85–89.9	
<b>B</b>	<b>3.000</b>	80–84.9	
<b>B–</b>	<b>2.667</b>	75–79.9	
<b>C+</b>	<b>2.333</b>	70–74.9	
<b>C</b>	<b>2.000</b>	65–69.9	
<b>C–</b>	<b>1.667</b>	60–64.9	
<b>F</b>	<b>0</b>	Below 60	

\* Source: <http://www.syr.edu/registrar/students/grades/faq.html>

### Schedule

Week	Topics	Readings	Activities and Dues
1, 1/27	Overview of the course Digital curation and the big picture <ul style="list-style-type: none"> <li>A view of cultural heritage</li> <li>A view of open science</li> <li>Terminology of digital curation</li> <li>The FAIR principles</li> </ul>	<b>Required:</b> Harvey & Oliver, chapters 1 & 2 <b>Skim through:</b> Cameron, chapter 1;	Case study: Tweets as digital cultural memory data
2, 2/3	The lifecycles <ul style="list-style-type: none"> <li>Research lifecycle</li> <li>Data lifecycle</li> <li>Digital curation lifecycle</li> <li>Datasets, data collections, data objects</li> <li>Data documentation</li> </ul>	<b>Required:</b> Harvey & Oliver, chapters 3 & 4 <b>Skim through:</b> Cameron, chapters 2 & 3	Lab 1: Data documentation
3, 2/10	Infrastructures for digital curation <ul style="list-style-type: none"> <li>Digital repositories, digital libraries, digital archives, and digital collections</li> <li>Requirements</li> <li>Technologies</li> </ul>	<b>Required:</b> Banerjee & Reese Jr., chapter 1 <b>Skim through:</b> Cameron, chapters 4 & 5	Lab 2: Collections as data
4, 2/17	Data repositories and workflows <ul style="list-style-type: none"> <li>Workflow types</li> <li>Repository platforms (<a href="https://research-data-network.readme.io/docs/repository-platforms">https://research-data-network.readme.io/docs/repository-platforms</a>)</li> </ul>	<b>Required:</b> Banerjee & Reese Jr., chapters 2 & 3 <b>Skim through:</b> Harvey & Oliver, chapters 5-7;	Case study 1: Decision on data repository solutions

## IST676 Digital Curation Spring 2022

	<ul style="list-style-type: none"> <li>Case study</li> </ul>	Cameron, chapters 6 & 7	
5, 2/24	Digital curation lifecycle in action: designing and creating data <ul style="list-style-type: none"> <li>Data format standards</li> <li>Metadata</li> <li>Curation-ready data</li> <li>Standards &amp; policies</li> <li>Structures of data</li> </ul>	<b>Required:</b> Harvey & Oliver, chapters 9 & 10 <b>Skim through:</b> Banerjee & Reese Jr., chapter 5	Lab 3: Design and creation of curation-friendly data
6, 3/3	Digital curation lifecycle in action: appraisal and data ingestion <ul style="list-style-type: none"> <li>What should be kept?</li> <li>For how long?</li> <li>Cost of keeping data</li> <li>Disposal of data</li> <li>Data ingestion workflows and tools</li> </ul>	<b>Required:</b> Harvey & Oliver, chapters 11 & 12 <b>Skim through:</b> <a href="#">NSF data management plan samples</a> (read a couple of samples)	Case study: Appraisal and selection policies
7 3/10	Digital curation lifecycle in action: preserving and storing data <ul style="list-style-type: none"> <li>Preservation, curation, archiving—what are the differences?</li> <li>Preservation methods</li> <li>Data migration</li> <li>Storage requirements</li> <li>Data security</li> <li>Storage solutions</li> </ul>	<b>Required:</b> Harvey & Oliver, chapters 13 & 14 <b>Skim through:</b> Lawrence et al., page 1-15.	Case study 2: Digital preservation
8, 3/17	Spring break. No class		
9, 3/24	Digital curation lifecycle in action: data access, use, and reuse <ul style="list-style-type: none"> <li>Discovery services</li> <li>Access control and authentication</li> <li>Legal issues</li> <li>Transform</li> </ul>	<b>Required:</b> Harvey & Oliver, chapter 15 <b>Skim through:</b> Banerjee & Reese Jr., chapter 9	Case study 3: Linking data for discovery
10, 3/31	Data services by libraries (Guest speaker: Dr. Renata Curty, Research Data Specialist, UC Santa Barbara Library) <ul style="list-style-type: none"> <li>Training</li> <li>Archiving for community and personal memories</li> <li>Consultation for data curation and preservation projects</li> <li>Developing data products</li> <li>Data publishing and sharing</li> </ul>	<b>Required:</b> <a href="#">Yoon &amp; Schultz, 2017</a> <b>Skim through:</b> <a href="#">Digital curation @ Smithsonian, LC Digital Strategy</a>	Lab 4: Creating open datasets  Assignment 1: Planning a digital data project
11, 4/7	Data products <ul style="list-style-type: none"> <li>What is a data product?</li> <li>Types of data products</li> <li>Data product design: Methods and tools</li> </ul>	<b>Required:</b> Patil, DJ, (2012). Data Jujitsu: the art of turning data into product.	

## IST676 Digital Curation Spring 2022

12, 4/14	Data Carpentry Workshop: OpenRefine Tutorial	<b>Required:</b> <a href="#">OpenRefine website</a>	Lab 5: OpenRefine exercise
13, 4/21	Data ethics framework <ul style="list-style-type: none"> <li>• Government, corporate, academic and cultural institutions, and individuals as digital data producers</li> <li>• Ownership of data</li> <li>• Ethical issues in data curation and services</li> <li>• Compliance to regulations</li> </ul>	<b>Required:</b> <a href="#">Data Ethics Framework</a> (U.S. Federal Data Strategy) <b>Skim through:</b> <a href="#">Data Ethics Framework</a> (GOV.UK)	Assignment 2: The biggest problem
14, 4/28	Wrap-up; Guest speaker: Wendy Kozlowski, Data Curation Specialist   RDMSG Coordinator, Cornell University Library <ul style="list-style-type: none"> <li>• Current development</li> <li>• Near future outlook</li> <li>• Discussion: how to prepare for a changing data-driving environment?</li> </ul>	To be decided.	
15, 5/5	<ul style="list-style-type: none"> <li>• Project presentation</li> </ul>		Final project due