FY 2019 Web Archiving Strategy
Yale University Web Archiving Working Group, May 2018 (membership updated 9.5.18)

Executive summary
This strategy builds upon the FY’17 Web Archiving Strategy, guided by the Web Archiving Working Group (WAWG) (See Appendix A) and implemented by a distributed group of web archiving practitioners. It recognizes that the greatest accomplishments to date have been the regular convening of web archiving professionals to coordinate efforts and exchange ideas and the development of resources for Web Archiving Practice and Policy considerations. It also recognizes that less progress was made towards FY’17 strategy in regard to community development and little progress was made towards a sustainable funding model.

In the next two years, the WAWG will coordinate efforts to mature shared web archiving practices across Yale. It will launch and lead a Web Archiving Community of Practice (CoP) program to include expert speakers, facilitated conversations, training workshops and use case demonstrations. It will continue to develop practice and policy guides and disseminate them through the CoP and existing University communication channels. Finally, it will aspire to use this grass roots approach to raise awareness among senior leaders and to gain entry to discuss priorities and sustainable investments needed to realize them.

Web Archiving Overview
The World Wide Web serves as a primary communication medium of the twenty-first century. Yet despite its ubiquity, web content is continually updated, replaced, deleted or lost, which threatens unique and valuable information. If no action is taken, important primary resources will be lost to the University and current and future scholars and students. Yale University needs to act now in order to protect these digital intellectual and cultural heritage assets and to comply with policy and contractual obligations for data retention. Preservation is fundamental to the University’s mission to improve the world today and for future generations through outstanding research and scholarship, education, preservation, and practice. In keeping with these priorities, the Web Archiving Working Group (WAWG) recommends the following principles and strategy to advance web archiving at Yale:

- Preservation and persistent access to web content is fundamental to the mission of the University across all of its functional domains
- Collaboration across Yale units, with colleagues at other institutions, with content creators and with content users is essential to successful web archiving practice
- Shared web archiving and access policies, procedures and practices are essential to an effective and coherent University approach
- Web archiving initiatives require sustainable commitment to staff and financial resources in order to succeed
- The rights and intentions of content creators are important and will be taken into account prior to harvesting web content

Objectives
The objectives outlined below formulate a two-year vision to support web archiving at Yale. They provide a framework to articulate stakeholder value, foster community and professional collaboration, secure funding, streamline and expand harvesting, and provide essential guidance on rights issues.

1. Stakeholder Value
   In FY’19 and FY’20, understand and document specific web archiving outcomes, opportunities and gaps in essential knowledge, tools and resources available to web archiving practitioners. In FY’20, understand senior stakeholder priorities and sustainable investments needed to realize them.

2. Collaboration
   In FY’19 and FY’20 web archiving practitioners will no longer practice in isolation. Practitioners will be connected to a Community of Practice that shares expertise, methods, policy understanding, skills and comradeship.

3. Funding
   In FY’19 the WAWG will continue to coordinate the federated approach to Yale’s web archiving effort through the marginal efforts of participating staff and request $15,000 from existing event and staff development budget lines to fund the CoP operational expenses such as speaker travel costs, training materials and refreshments. Prior to FY’20 the budget will be reviewed and reset according to priorities.

4. Practice
   In FY’19 the CoP, with guidance from the WAWG, will complete practice guides currently in process and develop new ones, as the need arises.

5. Policy, Procedures and Guidance
   In FY’19 the CoP, with guidance from the WAWG, will complete rights guides currently in process and develop new guidelines, as the need arises.

Strategy
Two years ago the Web Archiving Working Group developed an aggressive strategy to work collaboratively to develop initial practices, policies and procedures to advance web archiving on campus while simultaneously working with stakeholders to understand the web archiving value proposition and to build a campus-wide sustainable approach.

The group had considerable success with the former but was not able to significantly raise awareness among senior leadership as to the need for central investments to leverage and coordinate the existing departmental investments. At the same time, web archiving practitioners have found themselves short of the resources, skills and policies needed to advance web archiving in units not currently active in the practice.

Therefore, the WAWG is shifting its strategy to focus its efforts on a two-year program to develop and lead a Web Archiving Community of Practice (CoP). The goal of the CoP is to raise
awareness of the practice of web archiving at Yale and worldwide, to engage and support practitioners in meeting their web archiving goals, to provide educational opportunities and to develop shared practices through the exchange of case studies and the development and communication of practice guidelines, policies and procedures.

Over the next two years, the WAWG aspires to use this “grass roots” approach not only to advance web archiving at Yale but also to better understand the state of web archiving across the university in order to formulate appropriate, aligned and sustainable recommendations for the future.

Tactics
This section lists the key actions to be taken over the next two years for Yale to realize the goals set forth in this web archiving strategy (See Appendix B: Timeline).

Stakeholder and Community Development
1. Develop and lead a Web Archiving Community of Practice Program
   a. Develop the program framework and calendar of event types.
   b. Develop the communications plan.
   c. Develop recruitment plan and distribution list.
   d. Identify internal and external presenters.
   e. Identify community leadership for program.
   f. Deliver program.

Web Archiving Practice
2. Finalize recommendations on description and cataloging of archived sites. (FY’17, Goal 21)
3. Finalize recommendations for preservation of archived sites. (FY’17, Goal 22)

Resources
4. Secure FY’19 funding for CoP program
5. Secure FY’20 funding for CoP program

Policy, Procedures and Guidance
6. Approve the shared practice model for web archiving at Yale. (FY’17 Goal 17)
7. Finalize the comprehensive framework for web archiving rights. (FY’17 Goal 18)
8. Finalize the templates for statements of intent to archive and procedures for addressing ownership and rights. (FY’17, Goal 18)
Appendix A. Background and Web Archiving Working Group

Background
Although a few Yale staff members have been involved in web archiving initiatives with colleagues at other institutions, until recently there had been no concerted Yale effort to capture and preserve web content. In spring 2014, the Yale Center for British Art initiated a test of Archive-It, a web harvesting service provided by the Internet Archive, as a tool for capturing content from Yale websites and social media. After a successful test period, in 2015 the Beinecke Library, Yale ITS, Manuscripts and Archives, and the Yale Center for British Art entered into a one-year contract with Archive-It. The Archive-It service enables account holders to collect portions of the web and provide full text searching of captured URLs and metadata. Among the sites to be crawled this year are Yale University websites, select social media sites, and websites related to Beinecke’s and Manuscripts and Archives’ collecting areas.

This pilot year of web archiving has provided Yale with a chance to experiment with a new method of collection development, explore possibilities for describing and providing access to the captured content, and collaborate across campus units. The University-wide Web Archiving Working Group (WAWG) was established in 2015 to support web archiving activities at Yale and develop a web archiving strategy for the University.

Charge
Develop a web archiving strategy for Yale University, including website harvesting, description of the archived web content, development of access methods, and investigation and management of rights issues. The group will consult with other staff members as needed about digital preservation, description, and discovery; it will also maintain necessary contracts.

Current Members
Rachel Chatalbash (Co-Chair, Yale Center for British Art)
Alicia Detelich (Yale University Library, Manuscripts and Archives)
Daniel Dollar (Yale University Library) (Sponsor)
Jason Eiseman (Lillian Goldman Law Library)
Melissa Fournier (Yale Center for British Art)
Heather Gendron (Co-Chair, Robert B. Haas Family Arts Library)
Greta Graf (Digital Preservation Services Team, YUL)
Louis King (Information Technology Services)
Haruko Nakamura (International Collections and Research Support & Humanities Collections and Research Education)
Pam Patterson (Center for Teaching and Learning)
Alice Prael (Digital Accessioning Archivist, Yale Special Collections)
## Appendix B. Timeline

<table>
<thead>
<tr>
<th>Objectives</th>
<th>FY’19</th>
<th>FY’20</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stakeholder and Community Development</strong></td>
<td>• Develop and lead Web Archiving Community of Practice Program</td>
<td>• Lead Web Archiving Community of Practice Program</td>
</tr>
<tr>
<td><strong>Web Archiving Practice</strong></td>
<td>• Finalize recommendations on description and cataloging of archived sites. (FY’17, Goal 21)&lt;br&gt;• Finalize recommendations for preservation of archived sites. (FY’17, Goal 22)</td>
<td>• TBD by CoP in FY’19</td>
</tr>
<tr>
<td><strong>Resources</strong></td>
<td>• Secure FY’20 funding for CoP program</td>
<td>• Develop FY’21 Web Archiving Strategy&lt;br&gt;• Secure FY’21 funding</td>
</tr>
<tr>
<td><strong>Policy, Procedures and Guidance</strong></td>
<td>• Approve the shared practice model for web archiving at Yale. (FY’17 Goal 17)&lt;br&gt;• Finalize the comprehensive framework for web archiving rights. (FY’17 Goal 18)&lt;br&gt;• Finalize the templates for statements of intent to archive and procedures for addressing ownership and rights. (FY’17, Goal 18)</td>
<td>• TBD by CoP in FY’19</td>
</tr>
</tbody>
</table>