NDSA Web Archiving Survey

Introduction

In 2011 and 2013, the National Digital Stewardship Alliance (NDSA) conducted surveys of U.S. organizations currently or prospectively engaged in web archiving to better understand the landscape: similarities and differences in programmatic approaches, types of content being archived, tools and services being used, access modes being provided, and emerging best practices and challenges.

The resulting reports are available here:

http://www.digitalpreservation.gov/ndsa/working_groups/documents/ndsa_web_archiving_survey_report_2012.pdf

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_USWebArchivingSurvey_2013.pdf

The NDSA is releasing an updated survey in 2015 to continue to track the evolution of web archiving programs in the United States. The aggregate responses will be reported to NDSA members and summary results will be shared publicly.

Please allow 15 minutes to complete the survey. Thank you for your participation!

Next

Powered by SurveyMonkey®

NDSA Web Archiving Survey

About Your Organization

*** 1. Name of your organization**

2. Type of organization

- Archive
- Historical Society
- College or University
- Museum
- Public Library
- Commercial
- Consortium
- K12 School
- Government: Federal
- Government: State
- Government: Local
- Other (please specify)

3. Do you or your organization belong to any of these three groups? Check all that apply.

- [ ] International Internet Preservation Consortium (IIPC) http://netpreserve.org
- [ ] National Digital Stewardship Alliance (NDSA) http://digitalpreservation.gov/ndsa
- [ ] Society of American Archivists Web Archiving Roundtable http://www2.archivists.org/groups/web-archiving-roundtable

4. What is the status of your web archiving activity? Choose one.

- ( ) Planning / considering archiving but haven't started yet
- ( ) Pilot / testing
- ( ) Production / actively capturing
- ( ) Have collected content in the past but aren't currently collecting

Powered by SurveyMonkey®

See how easy it is to create a survey.

NDSA Web Archiving Survey

Archiving Program Information

5. How does the state of your organization's web archiving program compare to what it was two years ago? Choose one.

- [ ] Significant progress
- [ ] Some progress
- [ ] About the same
- [ ] Slightly worse off
- [ ] Much worse off
- [ ] Other or comment (use this box to provide an alternate answer or commentary on your answer above)

6. On what dimensions of the Web Archiving Lifecycle Model (https://archive-it.org/static/files/archiveit_life_cycle_model.pdf) has your organization made the most progress? Check three.

- [ ] Vision and Objectives
- [ ] Policy
- [ ] Resources and Workflow
- [ ] Risk Management
- [ ] Appraisal and Selection
- [ ] Scoping
- [ ] Data Capture
- [ ] Quality Assurance and Analysis
- [ ] Storage and Organization
- [ ] Preservation
- [ ] Metadata / Description
- [ ] Access / Use / Reuse

7. On what dimensions of the Web Archiving Lifecycle Model (https://archive-it.org/static/files/archiveit_life_cycle_model.pdf) has your organization made the least progress? Check three.

- [ ] Vision and Objectives
- [ ] Policy
- [ ] Resources and Workflow
- [ ] Risk Management
- [ ] Appraisal and Selection
- [ ] Scoping
- [ ] Data Capture
- [ ] Quality Assurance and Analysis
- [ ] Storage and Organization
- [ ] Preservation
- [ ] Metadata / Description
- [ ] Access / Use / Reuse

8. What are the goals of your web archiving activity? Check all that apply.

- [ ] Archive your own or affiliated web content as a type of institutional record (e.g., university archives archiving the university website or state library archiving state agency websites)
- [ ] Archive content from other organizations or individuals for future research (e.g., research library archiving third-party web content as part of topical collection building or special collections department archiving web content associated with a manuscript donor)
- [ ] Other (please specify)

9. What year did your organization begin archiving web content?

10. How much full-time equivalent (FTE) staff time does your organization dedicate to web archiving?

For example, if your organization's web archiving activity relied upon one dedicated, full-time employee and 50% of the time of three other full-time employees, you would indicate "more than 1, less than 3".

- ○ .25
- ○ .5
- ○ .75
- ○ 1
- ○ more than 1, less than 3
- ○ 3 or more

11. What are the top three considerations for the development of your web archiving program?

- ☐ Access and use (e.g., researcher interactions, web analytics, use cases)
- ☐ Cost (e.g., budgeting, service allowance utilization, staffing level requirements)
- ☐ Data volume (e.g., data volume collected, objects collected, acquisitions statistics)
- ☐ Institutional buy-in (e.g., programmatic growth, stakeholder testimonials, resource commitments)
- ☐ Loss (e.g., link and/or reference rot of archived resources)
- ☐ Quality (e.g., accuracy, completeness, replay fidelity)
- ☐ Risk management (e.g., permission responses, takedown requests, policy conformance)
- ☐ Other (please specify)

[ ]

12. What are the top three staff skills that are essential to the development and success of web archiving in your organization?

- [ ] Appraisal and selection (e.g., determining what web content to collect)
- [ ] Archiving tools (e.g., configuring or operating web archiving tools)
- [ ] Collaboration and communication (e.g., advocacy, coordination, marketing, or outreach)
- [ ] Domain expertise (e.g., knowledge of subjects that are the focus of web archiving)
- [ ] Metadata (e.g., familiarity with metadata standards, cataloging experience)
- [ ] Quality assurance (e.g., analyzing and troubleshooting web archive quality issues)
- [ ] Software development (e.g., able to develop software or web applications)
- [ ] Web technologies (e.g., familiarity with web architecture, design, formats, or platforms)
- [ ] Other (please specify)

13. What types of content do you have concerns about your capacity to archive? Check all that apply.

- [ ] Audio
- [ ] Blogs
- [ ] Databases
- [ ] Interactive media
- [ ] Social media
- [ ] Video
- [ ] Other (please specify)

14. In what areas of web archiving are you most interested in collaborating? Check all that apply.

- [ ] Best practices for policy and risk management
- [ ] Capture configuration and optimization
- [ ] Collaborative collection development
- [ ] Input on APIs and standards
- [ ] Metadata standards and application
- [ ] Quality assurance techniques and strategies
- [ ] Tool development, documentation, and user feedback
- [ ] Other (please specify)

15. What barriers do you face for collaborating on web archiving? Check all that apply.

- [ ] Still in planning / pilot stage; not much to share
- [ ] Lack of institutional support
- [ ] Lack of time to spend on collaborating
- [ ] Using a proprietary system
- [ ] Institutional policies prohibit or limit collaboration
- [ ] Other (please specify)

Powered by SurveyMonkey®

NDSA Web Archiving Survey

Tools and Service Providers

16. If capturing locally, what tool(s) do you use? Check all that apply.

- [ ] Adobe Web Capture
- [ ] Grab-a-Site
- [ ] Heritrix
- [ ] HTTrack
- [ ] Teleport Pro
- [ ] Web Archiving Integration Layer (WAIL)
- [ ] Web Curator Tool
- [ ] WebRecorder
- [ ] Wget
- [ ] Other (please specify)

17. If you are (or have been) using an external service for data capture, which one(s) do you use? Check all that apply.

- [ ] Aleph Archives
- [ ] Archive-It
- [ ] Hanzo Archives
- [ ] California Digital Library's Web Archiving Service
- [ ] Internet Archive's contract crawling services
- [ ] OCLC Web Harvester
- [ ] PageFreezer
- [ ] WebPreserver
- [ ] Other (please specify)

18. If you are (or have been) using an external service for data capture, have you replicated any of your data to another repository?

- ( ) Yes
- ( ) No

19. If so, to what kind of repository have you replicated the data? Check all that apply.

- [ ] Local repository
- [ ] External preservation service provider

20. If you have not replicated any of your data to another repository, why not? Check all that apply.

- [ ] Trust web archiving service provider
- [ ] Building local infrastructure
- [ ] No place to store / maintain it
- [ ] Not sure what we'd do with it once we got it
- [ ] Other (please specify)

Powered by SurveyMonkey®

See how easy it is to create a survey.

NDSA Web Archiving Survey

Access and Discovery

21. What kind(s) of access does your organization itself provide? Check all that apply.

- [ ] URL search
- [ ] Full-text search
- [ ] Browse list by URL
- [ ] Browse list by title
- [ ] Catalog records: collection-level description
- [ ] Catalog records: item-level description
- [ ] Finding aids
- [ ] Application programming interfaces (APIs)
- [ ] WARC files
- [ ] Derivative datasets
- [ ] Other (please specify)

22. Do you have active researchers utilizing your web archives?

- ( ) Yes
- ( ) No
- ( ) I don't know

23. If yes, could you provide a summary of how researchers are using your web archives?

5 / 7 71%

Prev Next

Powered by SurveyMonkey®

See how easy it is to create a survey.

NDSA Web Archiving Survey

Archiving Policies

24. Please indicate your typical approach regarding notifying and seeking permission from content owners when capturing and providing access to their web content. Choose one for each row.

	No action	Notifying	Requesting permission
Capturing	○	○	○
Providing restricted access	○	○	○
Providing public access	○	○	○

Other (please specify)

25. Is any of the web content archived by your organization embargoed as a matter of policy before being made accessible? Choose one.

- ○ Yes
- ○ No
- ○ Considering whether to implement embargo and/or what length
- ○ Not applicable (i.e., dark archive)

26. If you selected "Yes" for the previous question, how long is your organization's access embargo?

- ○ 6 months
- ○ 1 year
- ○ Other (please specify)

27. Does your organization have archiving policies specifically related to social media?

- ○ Yes
- ○ No

28. If so, would you be willing provide a link to or provide a summary of your approach?

29. Do you respect robots.txt when capturing?

- ○ Always
- ○ Never
- ○ Sometimes / it depends
- ○ Don’t know

30. If you selected "Sometimes / it depends" for the previous question, under what circumstances does your organization ignore robots.txt? Check all that apply.

- ☐ If your organization owns the copyright/has some other special access right (e.g., state archive collecting state agency websites)
- ☐ If permission is secured or appropriate notices have been sent
- ☐ In order to capture essential content (e.g., stylesheets, images, etc.)
- ☐ Other (please specify)

31. What resources has your organization relied upon in the development of its own copyright and access policies? Check all that apply.

- [ ] ARL Code of Best Practices in Fair Use for Academic and Research Libraries http://www.arl.org/focus-areas/copyright-ip/fair-use/code-of-best-practices
- [ ] Oakland Archive Policy http://web.archive.org/web/20140812200246/http://www2.sims.berkeley.edu/research/conferences/aps/removal-policy.html
- [ ] Section 108 Study Group Report http://www.section108.gov/
- [ ] Consultation with legal counsel
- [ ] Statutory authority
- [ ] Web archiving policies and practices of other organizations
- [ ] Previous NDSA Web Archiving Survey reports
- [ ] Other (please specify)

If you are willing to share your web archiving policies, please send a copy or provide a link to ndsa@loc.gov.

Powered by SurveyMonkey®

See how easy it is to create a survey.

NDSA Web Archiving Survey

Thank you for participating in the 2015 NDSA Web Archiving Survey!

Results will be shared in 2016.

7 / 7 100%

Prev Done

Powered by

SurveyMonkey®

See how easy it is to create a survey.