2014

Florida Digital Action Plan: Statewide Digitization Survey Report

Prepared by Tom Clareson, LYRASIS; Liz Bishoff, The Bishoff Group; and Jason Kucsma, METRO, for the Florida Department of State, Division of Library and Information Services.
April 18, 2014

Florida Digital Action Plan: Statewide Digitization Survey Report

Executive Summary

In October 2013 work began on a *Florida Statewide Digital Action Plan* project to gather data on current and planned digital activities among the state's cultural heritage institutions; build understanding about digitization issues across the cultural community; draft policies and best practices; and develop technical strategies to support a statewide digitization program. Central to the activities of the Steering Committee and Work Groups is the question: "How do Florida's cultural heritage organizations create a single point of discovery of Florida's digital collections for residents?"

As part of the initial data gathering step of the project, top administrators and professionals at a wide variety of cultural heritage institutions participated in the *Florida Statewide Digitization Survey* in early 2014. Survey results showed that digitization activities were underway at many Florida institutions with a wide range of staff and budget sizes.

The *Florida Statewide Digital Survey*, conducted between January 7 and February 17, 2014, had a total of 101 responses. The survey was sent to Florida cultural heritage institutions, including: the Florida Association of Museums (220 members); the Florida Department of State's Public and Special Libraries email list (280 organizations, including 153 public libraries and 127 special libraries); the Society of Florida Archivists (13 active institutions); 39 state colleges and universities; and 31 independent colleges and university libraries in Florida, for a total of 583 institutions. The response rate for the survey was approximately 17%. Survey participation from the public and academic library and archives sectors was particularly strong.

Key findings:

- Of the 101 survey respondents, 71 (71%) are creating and/or acquiring digital collections, and 8 (8%) plan to create/acquire digital collections in the next three years. These 71 organizations reported having 13.8 million metadata[1] records, providing access to collections across a range of topics and types of materials.
- Funding of digital collections is largely supported by the organizations' operating budget; however, 21% (15) of the 71 respondents creating and/or acquiring digital collections indicated that they did not have a specific budget source for digital collection activities. The majority of the respondents' digital program budgets were below $50,000.
- 33% (13 of 39 respondents) of Florida cultural heritage organizations have 75-100% of their collections available to the public, while only 10.3% (4) do not have their collections available.

---

[1] A total of 13,862,310 metadata records were reported by 39 institutions. One respondent reported having 11 million metadata records, while 3 reported having between 300,000 and 1,000,000+ metadata records. Others reported as low as 15, while many reported between 1,000 and 10,000.

April 18, 2014

- Many respondents expect their collections to grow over the next three years. Currently, 31% (20) of the 65 respondents have less than 1 terabyte (TB) of their master files in digital format, while 8.6% (6) have more than 11TB. By 2017, 25 predict they will have up to 5TB, and 3 predict they will have 50-501+TB.
- 50.7% (34 of 67 respondents) of organizations have a digital collection management system. The majority of the 49.3% (33) that do not are museums and historical societies.
- 23% (9 of 39 respondents) support Open Archival Information Protocol Metadata Harvesting (OAI-PMH), while 17.9% (7) support File Transfer Protocol (FTP), and 23.1% (9) support HTML data export. These data transfer mechanisms are key to creating shared repositories.

Recommendations: Florida's cultural heritage organizations are making a significant body of digital content on a wide range of topics available to Florida's residents. Several challenges exist moving forward; the recommendations below address these challenges:

- Expanding participation: Expanding the range of cultural heritage organizations that are engaged in creating and making their digital collections available to Floridians should be a key focus of any future efforts. Bringing in more cultural heritage organizations will create a wider depth and breadth of content, as well as diversify the awareness and use of the collections.
- Public access to digital collections: The methods by which organizations can provide their collections to the public must be expanded. Developing a single point of access to Florida's digital collections will help meet this goal.
- Digital Asset Management Systems: Just 50% of the organizations creating digital content have digital asset management systems that support access to their collections. Developing solutions that will expand the use of digital asset management systems will allow organizations to improve the local management of collections, as well as support the collaborative sharing of collections, both nationally and internationally.
- Training: The survey identified a lack of a programmatic approaches (which could be implemented on a regional or statewide basis, or by type of organization, such as museum/archival/library) that support small and medium sized organizations in their efforts to digitize their collections, or build digital collections through acquisition. Programmatic elements include: training/education in digital standards and best practices; platforms for access to digital content; digital preservation; and promotion of digital collections.

The consultants wish to thank the administration and staff of the Florida Department of State's Divisions of Library and Information Services, Cultural Affairs and Historical Resources, and the members of the *Florida Statewide Digital Action Plan* Steering Committee, for their insights and leadership throughout the survey project.

Introduction

In early 2014, the Florida Department of State's Divisions of Library and Information Services, Cultural Affairs and Historical Resources, working with the *Florida Statewide Digital Action Plan* Steering Committee, conducted a statewide digitization survey. Reaching out to museums, historical societies, archives and libraries, the survey examined: institutional demographic information; digital collection management; selection and acquisition issues; Florida collections; access to digital collections; information technology infrastructure; rights issues; collaboration and partnership issues; and digital preservation.
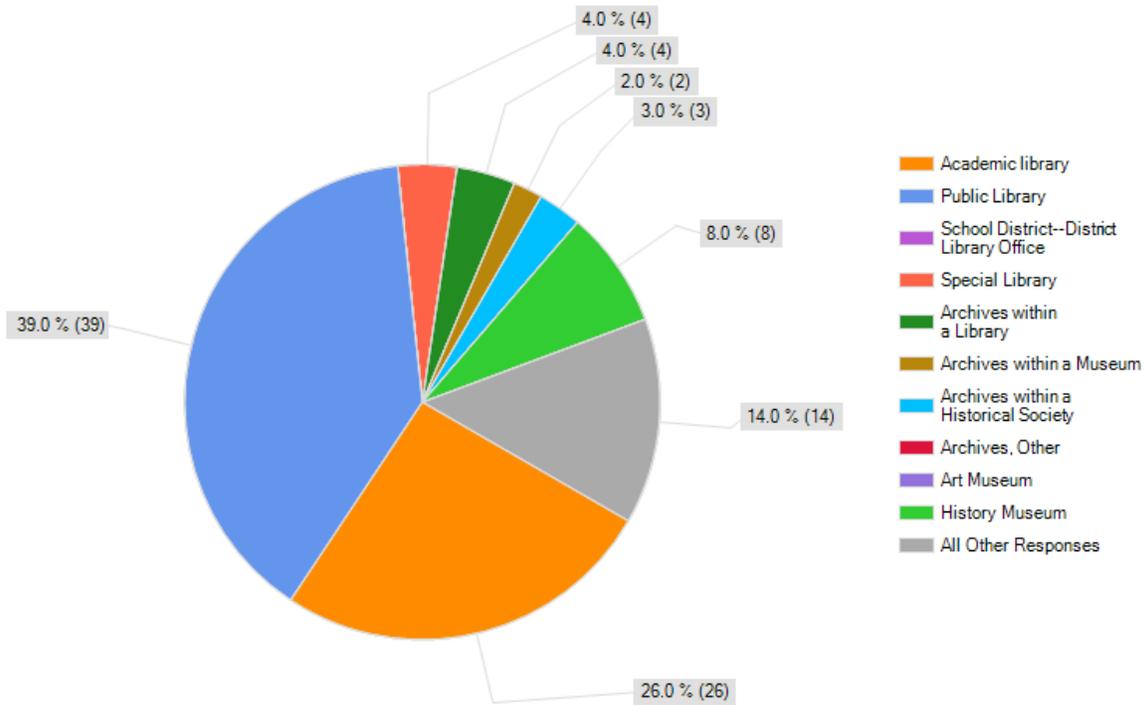
The results of the survey will provide information for a digital action plan supporting access to content available in Florida's cultural heritage organizations.

Survey Demographics

The Florida Statewide Digitization Survey, conducted between January 7 and February 17, 2014, had a total of 101 responses. The survey was sent to Florida cultural heritage institutions, including: the Florida Association of Museums (220 members); the Florida Department of State's Public and Special Libraries email list (280 organizations, including 153 public libraries and 127 special libraries); the Society of Florida Archivists (13 active institutions); 39 state colleges and universities; and 31 independent colleges and university libraries in Florida, for a total of 583 institutions. The response rate for the survey was approximately 17%.

The institution type most widely represented was public libraries with 39 responses (39% of the total), followed by academic libraries with 26 (26%) responses. 13 (13%) responses were museums: 8 history museums, 1 science museum, and 4 categorizing themselves as, "other museum types." 9 (9%) responses were archives: 4 within a library, 3 within a historical society and 2 within a museum. 4 (4%) responses were special libraries; 2 (2%) were historical societies, and 1(1%) was a school district/district library office. Seven institutions categorized themselves as "Other," and comprise: 3 government libraries and archives; 1 news library; 1 natural history museum; 1 museum and rare books library; and 1 special collections library.

April 18, 2014

What type of institution do you represent? (e.g., academic library, art museum or archive)

4.0 % (4)
4.0 % (4)
2.0 % (2)
3.0 % (3)
39.0 % (39)
8.0 % (8)
14.0 % (14)
26.0 % (26)

- Academic library
- Public Library
- School District--District Library Office
- Special Library
- Archives within a Library
- Archives within a Museum
- Archives within a Historical Society
- Archives, Other
- Art Museum
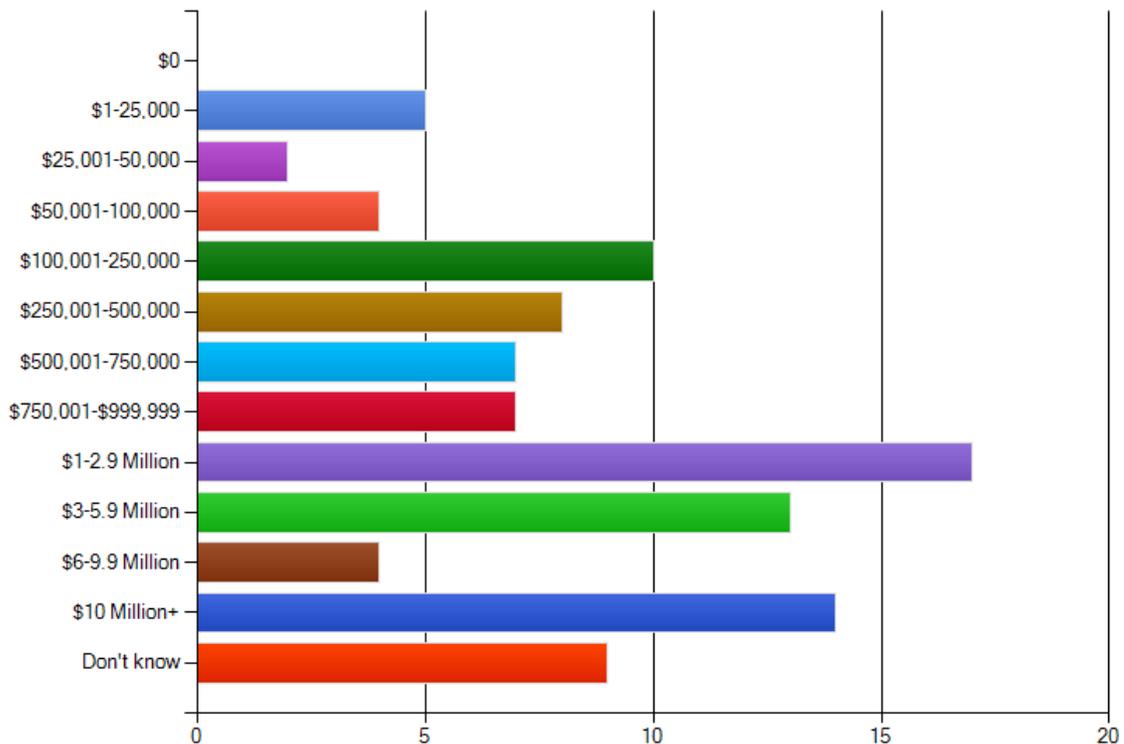- History Museum
- All Other Responses

In the demographic section of the survey, 100 of the 101 responding institutions provided data. The majority of the respondents (58, or 58% across all institution types) had Administrator (Dean or Director) as their primary role. Other survey respondents were: Librarians (13); Archivists (8); Curators (5); Digital Librarians (4); Registrar (3); and Information Technologists (2). Eight respondents categorized themselves as "Other," including: 4 Digital Collections Managers; 1 Historian; 1 Content Licensing Specialist; 1 Publications Specialist; and 1 Librarian.

When asked about the number of Full-time Equivalent (FTE) staff employed at the responding institutions, there was a wide variety of staffing levels. While the largest group of respondents (17) had 3 to 5 FTE staff, there were also 17 institutions with 101+ staff. Thirteen reported 11 to 20 staff; the same number reported 31 to 50 staff. There were 10 organizations with 6 to 10 staff. Six respondents, 3 from archival institutions, indicated they had 0 FTE. There were 7 respondents with 1 to 2 staff; 6 with 21 to 30 staff; 5 with 51 to 75 staff, and 4 with 76 to 100 staff. Two respondents did not know the number of FTE staff at their institution.

Survey respondents were asked to provide information on their institution's annual operating budget, including: staff; collection budgets; technology; facilities management; etc. As with staffing, there were a wide range of responses. The majority (17 organizations) reported annual operating budgets of $1,000,000 to $2,999,999. 14 reported budgets over $10 million; 13 reported budgets of $3,000,000 to

April 18, 2014

$5,999,999; 10 reported budgets of $100,001 to $250,000; 8 reported budgets of $250,001 to 500,000; 7 reported budgets of $500,001 to $750,000; 7 reported budgets of $750,001 to $999,999; 5 reported budgets of $1 to $25,000; 4 reported budgets of $50,001 to $100,000; 4 reported budgets of $6,000,000 to $9,999,999; 2 reported budgets of $25,001 to $50,000; and 9 institutions did not know their annual operating budget.

What is your institution's annual operating budget (including as appropriate, staff, collection budget, technology, facilities management, etc.)? For a definition of institution see Section A above.

Budget	Count
$0	0
$1-25,000	5
$25,001-50,000	2
$50,001-100,000	4
$100,001-250,000	10
$250,001-500,000	8
$500,001-750,000	7
$750,001-$999,999	7
$1-2.9 Million	17
$3-5.9 Million	13
$6-9.9 Million	4
$10 Million+	14
Don't know	9

Digital Collection Policies, Budgets and Staff

71 (71%) of respondents, including all of the archives responding to the survey, have digital collections. Of the 29 that did not (mostly from the public library sector), 8 replied that they would begin creating digital collections in the next three years, and 20 (16 of which are public libraries) said they would not.

Many of the organizations began creating digital collections in the past decade or before. 44 organizations began in 2001 or after, and 21 began before 2000. The largest number of respondents, 16 (22.5%) of 65, began digital collection creation between 2001 and 2005; 14 began between 2011 and 2014; 14 began between 2006 and 2010; 12 began between 1996 and 2000; 7 began between 1991 and 1995; 2 began prior to 1990; 2 did not know when they began; and 4 said their institution does not create digital collections.

April 18, 2014

When asked the year that their institution began collecting or acquiring digital collections, many of the date ranges were the same, but 16 (22.5%) of the 71 respondents said their institutions does not collect or acquire digital collections. Of the 51 that do, most (16) began between 2001 and 2005; 14 between 2011 and 2014; 8 between 2006 and 2010; 9 between 1996 and 2000; 3 between 1991 and 1995; 1 prior to 1990, and 3 did not know the year they began.

Respondents were asked if they had written policy documents addressing digital collections in a variety of areas, including:

- Mission and Goals
- Collection Development
- Emergency Preparedness
- Exhibits
- Preservation
- Strategic Planning
- Rights and Licensing

In all cases, the majority of respondents did not have these policies. The areas where some organizations did have policies addressing digital collections are: 25 in collection development; 22 in mission and goals; and 21 in rights and licensing (almost half of which are academic libraries). Institutions with larger budgets and more FTE staff more often developed these policies. 14 institutions had Rights and Licensing policies under development; 15 had digital Preservation policies under development; and 13 are currently developing Missions and Goals addressing digital collections. Policy development activity was especially prevalent among the archival respondents.

April 18, 2014

Does your institution have a written POLICY DOCUMENT that addresses digital collections in each of the following cases?
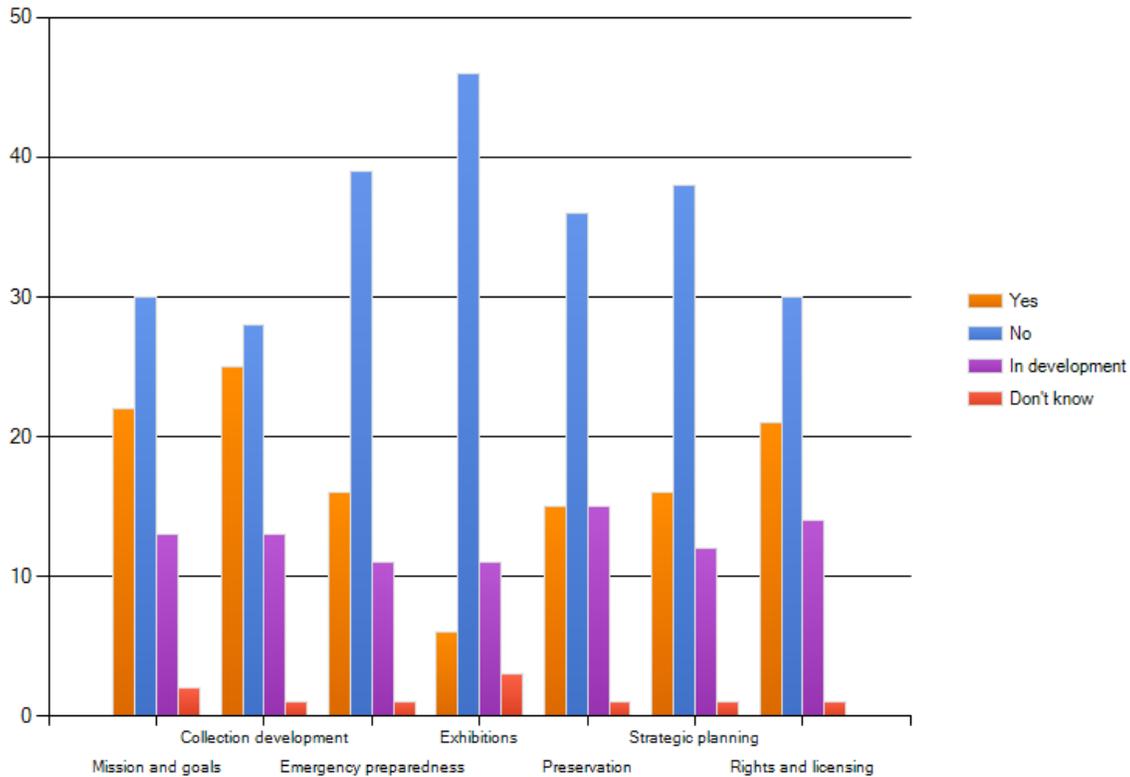
When asked for the number of Full-Time Equivalents (FTE), including staff and volunteers supporting their digital initiative, the majority, 39 institutions (56.5%) of 69 respondents, across all institution types and overall staff sizes, said between 1 and 3 FTE. 11 organizations reported no staff or volunteers supporting digital collections; 8 reported between 3 and 5; and 3 (mostly from organizations with larger annual operating budgets) reported between 6 and 10 FTE, between 10 and 15 FTE or 16+ FTE.

Looking at operating budgets for the digital collection initiative, including staffing, technology, etc., the largest number of respondents, 15 (21.7%) of 69, across all staff sizes, and mostly in public and special libraries, indicated no funding. 7 reported between $2,501 and $5,000; 6 reported between $1 and $2,500; 5 organizations each reported between $50,001 and $100,000, and between $250,001 and $500,000; 3 reported budgets of $500,001 and above; and 10 said they did not know.

By far the largest number of respondents, 53 (81.5%) of 65, said their organization's operating budget was a source of funding for their digital collections initiative. In this question, where organizations could indicate multiple answers for funding sources, responses included: grants, 23 (35.4%); fundraising, 11 (16.9%); sales of products associated with digital collections, 4; fees from activities, 2; funding through consortia, 2; endowment funding and user fees, 1; and 2 organizations claimed no budget source for a digital collections initiative. Of particular interest was that institutions with larger budgets were among the largest users of grant funding for digital collections activities.

April 18, 2014

<u>Digital Collection Development</u>

When organizations were asked why they are creating and/or acquiring digital collections, they provided a wide variety of answers. The 65 respondents could choose multiple reasons; including:

- To provide online access to materials: 60 (93.8%)
- To increase access to the collections: 57 (89.1%)
- To preserve the original by reducing handling: 53 (82.8%)
- For study and use by local users: 52 (81.3%)
- For study and use by remote users: 46 (71.9%)
- For marketing and promotion of the institution: 36 (56.3%)
- For identification (i.e., collections management, cataloging, etc.): 27 (42.2%)
- For publication: 24 (37.5%)
- To replace collection items in the event of a disaster, deterioration, etc.: 21 (32.8%)
- To contribute to a collaboration/consortium: 20 (31.3%)
- For documentation of condition: 11 (17.2%)
- Legally mandated to acquire materials: 6 (9.4%)
- To generate revenue: 6 (9.4%)

There are also a wide variety of criteria listed as being used to select materials for digitization. Again, this question allowed multiple answers by the same institution; percentages are based on a total of 64 respondents:

- There is a strong local interest in our materials and/or collections: 49 (76.6%)
- Materials are fragile or deteriorating: 42 (65.6%)
- Materials are of high value and digitizing will increase access: 40 (62.5%)
- Materials are heavily used: 31 (48.4%)
- There is a grant available to digitize materials in the collection: 16 (25%)
- Donor requested that the collection be digitized: 11 (17.2%)
- Digitized materials as a collection record only, and don't make the images publicly available: 5 (7.8%)
- Digitized everything in our collection: 4 (6.3%)
- Legally required to make the materials available online: 2 (3.1%)

Respondents were asked if their organization created digital collections from physical source materials, and 59 (89.4%, including all archival respondents) said yes, while 7 (10.6%) said no. The 59 respondents that created digital collections used a wide variety of material types as sources, including:

- Photographic prints: 54 (91.5%)
- Flat works on paper: 44 (74.6%)
- Text, manuscripts or other multi-page items: 37 (62.7%)
- Maps, architectural drawings and posters: 36 (61%)
- Film, both film negatives and glass plate negatives: 32 (54.2%)

8

- Books: 31 (52.5%)
- Newspapers: 29 (49.2%)
- Manuscripts (letters): 25 (42.4%)
- Two-dimensional works of art: 23 (39%)
- Video: 22 (37.3%)
- Three-dimensional objects: 19 (32.2%)
- Analog audio: 19 (32.2%)
- Three-dimensional artifacts: 18 (30.5%)
- Three-dimensional works of art: 17 (or 28.5%)
- Microfilm (10 or 16.9%)

Forty (60.6%) of the 66 responding institutions – the majority from the library sector, as well as all but one of the participating archival organizations -- are creating/acquiring born-digital collections. Eighteen (27%) are not and 8 (12.1%) organizations did not know. A wide variety of born-digital material types are being created or acquired.

The top born-digital formats being created include:

- Documents (PDFs, Word documents, spreadsheets, etc.): 32 institutions
- Photography or other still images: 26 institutions
- Digital video: 22 institutions
- Digital audio: 21 institutions
- Documentation or research data: 13 institutions
- Social media (blogs, websites, listservs, mailing lists, etc.): 12 institutions
- Maps: 12 institutions
- Non-licensed e-books and e-journals: 9 institutions
- Art or visual materials with database or digital components: 7 institutions
- Geospatial data: 5 institutions
- Application, operation systems or other software: 5 institutions
- Other numeric data sets: 3 institutions

Top types of born-digital materials that institutions are acquiring include:

- Photography or other still images: 26 institutions
- Documents (PDFs, Word documents, spreadsheets, etc.): 25 institutions
- Digital video: 24 institutions
- Digital audio: 18 institutions
- Documentation or research data: 11 institutions
- Non-licensed e-books and e-journals: 9 institutions
- Maps: 8 institutions
- Application, operating systems or other software: 6 institutions
- Social media (blogs, websites, listservs, mailing lists, etc.): 5 institutions

April 18, 2014

- Art or visual materials with database or digital components: 5 institutions
- Geospatial data: 4 institutions
- Other numeric datasets: 3 institutions

Digital Standards, File Formats and Collection Size

Respondents were asked to list the standards or best practices their institutions are using for digital content creation (digitization). Note that 64 to 66 respondents provided answers to the questions described in this section.

Please list the standards or best practices you are using for digital content creation (digitization). Select all that apply.
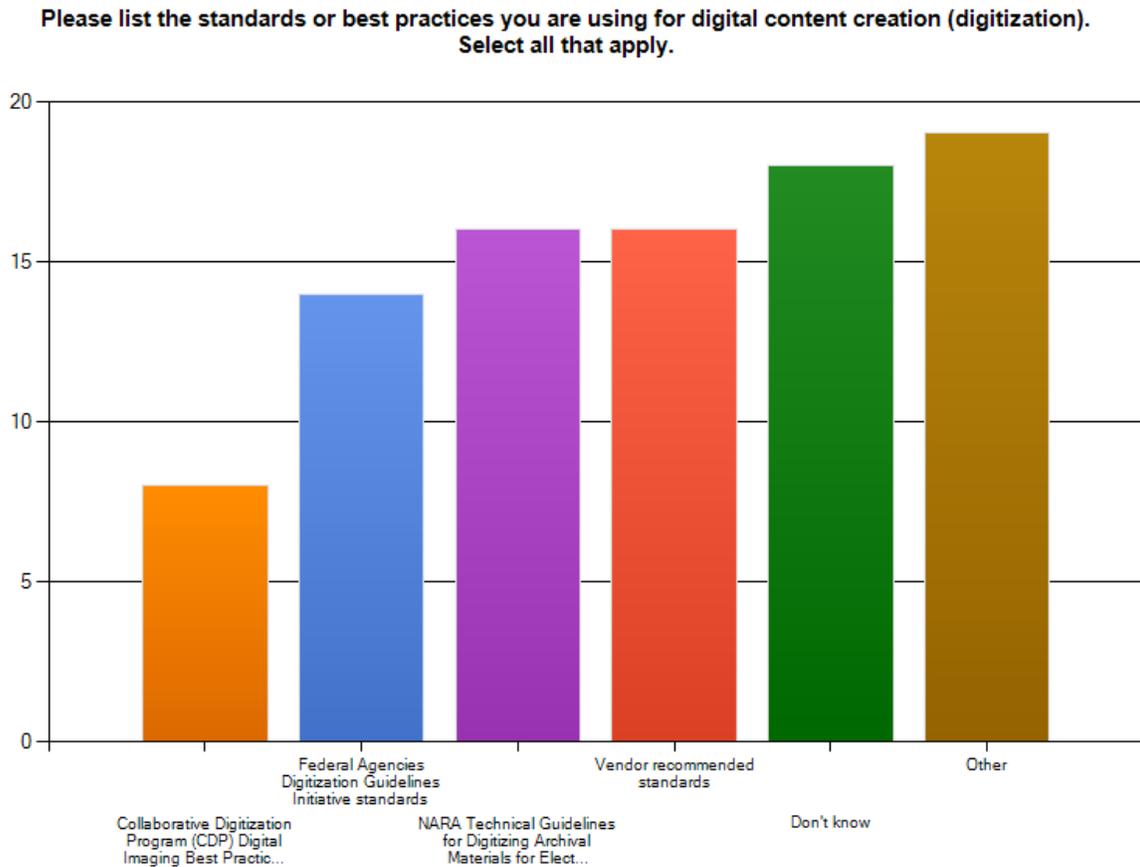
The majority of organizations, 19 (28.8%), did not choose the national standards listed in the answer categories for this question, and instead answered "Other." Standards and practices listed in these responses included, in-house or local standards at 5 institutions, no standards at 4 institutions and multiple standards at 4 others.

Eighteen (27.3%) organizations did not know the standards and best practices being used, and 16 (25.8%) are utilizing vendor recommended standards.

Sixteen organizations, mostly academic libraries, are using the, *NARA Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images.* Fourteen,

April 18, 2014

almost all from institutions with larger staff sizes, use, *The Federal Agencies Digitization Guidelines Initiative Standards (FADGI)*. Eight are utilizing, *The Collaborative Digitization Program (CDP) Digital Imaging Best Practices, Volume 2.*

Most organizations, 38 (59.4%), across all budget sizes, indicated they are creating digital collections in-house; 19 (29.7%) are creating collections both in-house, and by outsourcing to a vendor; and 7 (10.9%) are using outsourcing alone.

Institutions have adopted a variety of digital file formats for the creation of master files, whether created by their institutions or acquired. Top formats include:

- TIFF: 47 (72.3%)
- PDF/PDF-A: 32 (49.2%)
- JPEG: 31 (47.7%)
- WAV: 18 (27.7%)
- JPEG 2000: 9 (13.8%)
- AVI: 9 (13.8%)
- Don't know: 6 (9.2%)

Fifteen organizations listed "Other" for the file formats used; answers included: RAW files: 2 institutions; MP4: 2 institutions; MP3: 2 institutions; as well as single mentions of a variety of video-related file formats.

Respondents provided information on the storage size for their digital collections in 2014, reporting only master files, not derivatives. In general, the size of the digital collections are small, with 16 (31%) of the respondents having less than 1TB of data. 6 (8.6%) of the respondents have more than 11TB of data. The majority of organizations, 17 (26.2%), across all institution types, and especially in organizations with smaller staff and budget sizes, did not know the storage size for digital collections. Of those that did, the top answers were:

- 0-200GB: 12 institutions (18.5%) of the 65 total responding institutions
- 201-500GB: 4 (6.2%)
- 501-999GB: 4 (6.2%)
- 1-2TB: 11 (16.4%)
- 3-5TB: 5 (7.7%)
- 6-10TB: 6 (9.2%)
- 11-15TB: 2 (3.1%)
- 16-25TB: 1 (1.5%)
- 26-50TB: 1 (1.5%)
- 51+TB: 2 (3.1%)

April 18, 2014

Next, respondents were asked to estimate how large they expect storage to be for their digital collections in 2017. Again, the largest number of respondents, 22 (33.8%) of 65 total answering institutions, did not know. Other estimates included:

- 0-200GB: 7 institutions (10.8%)
- 201-500GB: 5 (8.2%)
- 501-999GB: 2 (3.3%)
- 1-2TB: 3 (4.9%)
- 3-5TB: 6 (9.8%)
- 6-10TB: 3 (4.9%)
- 11-15TB: 1 (1.6%)
- 16-25TB: 5 (8.2%)
- 26-50TB: 3 (4.9%)
- 51-100TB: 1 (1.6%)
- 101-500TB: 1 (1.6%)
- 501+ TB: 1 (1.6%)

Almost all institution types moved up at least one unit of measure of digital collection size between their 2014 figures and 2017 estimates.

Florida Collections

Survey respondents were asked about Florida related collections held by Florida libraries and cultural institutions that have been digitized, or are in digital formats, in a variety of subject matter areas. For all but one subject area, photographs were the leading format of materials that had been digitized. Geology was the only area where five respondents digitized mostly textual material.

When asked about the collections that could be digitized in the future, respondents said they wanted to digitize all formats of material (photographs; maps; manuscripts; audio; videos; books; texts; and other formats) in 28 of the suggested subject areas. Of all the subject areas, Local History (29 respondents), and Florida History (26 respondents), were the most popular for future digitization. A list of subject areas where institutions are interested in digitizing collections in the future appears as Appendix A.

Access to Digital Collections

When asked if they have a digital asset management system, 34 (50.7%) organizations, including 14 academic libraries and organizations with larger staffs, said yes; 33 (49.3%) said no.

The majority of the data on access to digital collections is based on 39 respondents that replied to questions in this section. The top two methods of providing public access to digital collections (again, respondents were able to provide multiple answers to this question) were through a website associated with the institution: 30 (76.9%), and through a digital asset management system, such as *CONTENTdm* or *DigiTool:* 29 (74.4%). Other responses included: through a website associated with another

April 18, 2014

institution: 15 (38.5%); via social media (*YouTube, Flickr, Pinterest*): 16 (41%); CD/DVD/Other removable media: 10 (25.6%); and standalone computers: 9 (23.1%).

The survey asked: "For all digital items, what percentage have an access copy available online?" Thirteen (33.3%) organizations, mostly from the library sector, reported 76-100% of their digital items have access copies available online; 9 (23.1%) have 1-25% available; 5 (14.3%) have 51-75%, with an access copy available online; 3 have 26-50%, with an access copy available online; 4 (10.3) do not have any of their digital items available online; 3 do not have any of their digital items available online, but they plan to in the next one to two years; 1 does not know the percentage; and 1 has not yet gone live with their digital content management system.

Of all the digital items, what percentage have an access copy available online?
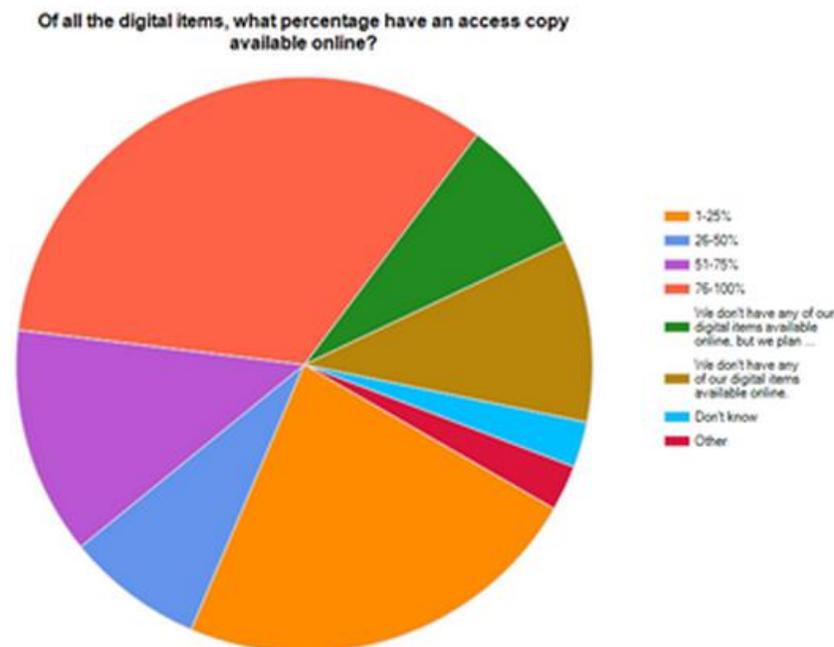
Digital items require metadata for discovery, access, management and preservation. When asked about the types of metadata created for their digital items, respondents answered:

- Descriptive metadata: 34 (87.2%)
- Administrative metadata: 26 (66.7%)
- Technical metadata: 23 (59%)
- Structural metadata: 18 (46.2%)
- Preservation metadata: 14 (35.9%)
- None: 4 (10.3%)
- Don't know: 1 (2.6%)

Academic libraries, and archives within a library, were the leading respondents for creation of all metadata types.

April 18, 2014

As a follow-up, organizations were asked what percentages of their digital items have metadata. Twenty-one (53.8%) institutions, mostly academic libraries, said between 76% and 100%; 7 (17.9) said between 51% and 75%; 5 organizations said none; 2 organizations said between 1% and 25%, or between 26% and 50%; and 2 did not know.

Respondents were asked how many metadata records have been created for their institution's digital collections. A total of 13,862,310 metadata records were reported by 39 institutions. One respondent reported having 11 million, while 3 reported having 300,000-1,000,000+. Others reported as low as 15, while many reported between 1,000 and 10,000. Some organizations, mostly from public, academic and special libraries, and history museum sectors, did not report, or did not know how many metadata records were in their collection.

Cataloging standards and controlled vocabularies used in preparing descriptive metadata includes:

- Library of Congress Subject Headings (LCSH): 22 (56.4%, primarily from academic libraries)
- Anglo-American Cataloging Rules - 2(AACR-2)/Resource Descriptive and Access (RDA): 20 (51.3%)
- Art and Architecture Thesaurus (AAT): 14 (35.9%)
- Local Rules: 12 (30.8%)
- Describing Archives – A Content Standard (DACS): 13 (33.3%)
- Thesaurus of Graphic Materials I and II (TGM I & II): 12 (30.8%)
- Chenhall Nomenclature for Museum Cataloging: 6 (15.4%)
- None: 4 (10.3%)
- Don't know: 4 (10.3%)
- Cataloging Cultural Objects (CCO): 3 (7.7%)
- Other: 12 (30.8%). Four of these are local, or in-house, standards or vocabularies.

Use of all of these cataloging standards, or controlled vocabularies, was most prevalent in institutions with the largest staff and budget sizes.

Metadata schemas being used for digitization work include:

- Dublin Core: 22 (56.4%, primarily academic libraries)
- MARC: 15 (38.5%)
- Metadata Object Descriptive Schema (MODS): 11 (28.2%)
- Encoded Archival Description (EAD): 10 (25.6%)
- Metadata Encoding Transmission Standard (METS): 9 (23.1%)
- None: 9 (23.1%)
- Preservation Metadata: Implementation Strategies (PREMIS): 4 (10.3%)
- NISO Draft Standard: Data Dictionary – Technical Metadata for Digital Still Images: 3 (7.7%)
- Don't know: 3 (7.7%)
- Visual Resources Association Core (VRA Core): 2 (5.1%)
- Public Broadcasting Core (PB Core): 2 (5.1%)

April 18, 2014

- Darwin Core: 2 (5.1%)

Eight institutions listed other metadata schemas they are using, including a combination of the schemas listed above, or locally developed schemas.

Digital Asset Management systems can be used to manage the full life cycle of digital objects, including: management of data creation; metadata repository; image repository, or linkage to the image repository; registry of preservation metadata; and as a means of providing access to users. The top types of systems used by survey respondents included:

- *OCLC's CONTENTdm*: 11 (28.2%)
- *Islandora*: 8 (20.5%, all academic libraries)
- Locally developed systems: 7 (17.9%)
- *PastPerfect*: 6 (15.4%)
- *SobeK*: 5 (12.8%)
- *Omeka*: 4 (10.3%)
- *ExLibris DigiTool*: 4 (10.3%)
- *D-Space*: 3 (7.7%)
- *Fedora*: 1 (2.6%)
- "Other" systems, including *Bepress Digital Commons*: 4, and Archon: 2

The largest institutions, defined by staff and budget size, most often implemented Digital Asset Management systems.

Metadata harvesting strategies supported by responding institutions (again, most often in organizations with larger budget and staff sizes) included:

- Don't know: 10 (25.6%)
- Open Archival Information Protocol for Metadata Harvesting (OAI-PMH): 9 (23.1%, mostly academic libraries)
- HTML (data export): 9 (23.1%)
- File Transfer Protocol (FTP): 7 (17.9%)
- Z39.50: 8 (20.5%)
- System supports OAI-PMH, but institution has not implemented it: 6 (15.4%)
- System supports FTP, but institution has not implemented it: 6 (15.4%)
- Organizations does not support any metadata harvesting capability: 3 (7.7%)
- System supports Z39.50, but institutions has not implemented it: 2 (5.1%)

Information Technology

Thirty-nine organizations provided data in this section, where respondents were able to provide multiple answers according to their level of activity. Responding institutions most often have their IT needs met through an internal IT department: 28 (71.8%); 12 (30.8%) outsource IT needs to an external

organization or vendor; 8 (20.5%) outsource to an internal IT unit other than the library/museum/cultural heritage institution; and 4 do not have an IT department.

IT services for digital collection management that are supported by responding institutions included: digital imaging (scanning, direct digital capture and digital photography) at 35 institutions; a collection management system/digital asset management system at 30 institutions; Web development/design at 29; an institutional repository at 18; and a preservation repository at 11.

<u>Rights Management</u>

Institutions were asked to rate the accuracy of a number of statements regarding rights management for digital materials. 32 institutions (82.1% of the 39 providing data) said it was very accurate, and that they, "Consider copyright and/or intellectual property concerns in managing digital collections." 15 organizations (40.5%) said it is very accurate, and that they, "feel confident making copyright licensing and digital copyright decisions about (their) digital collections." 12 organizations, including all archives within libraries, said it was not at all accurate, and that, "copyright and licensing concerns deter us from creating and preserving digital collections." 10 also said it was not accurate, and that they, "record and maintain rights metadata to limit delivery of collections to authorized users," although 7 respondents said that statement was very accurate. Finally, 19, a majority of respondents, said they ranked a 4 or 5 on a 5 point scale that the statement, "When we acquire collections, we acquire digital preservation rights," was very accurate.

More than half of the respondents, 22 (56.4%), have not updated their Deed of Gift/Donor Agreement to include digital collections; 13 organizations (33.3%, mostly academic libraries) have done this update; and 4 (10.3%) organizations did not know.

<u>Collaborative Activities</u>

Over half of the institutions participating in the questions about collaborative digital activities, 22 (56.4%, including almost all of the academic libraries taking part in the survey), said they collaborate with other libraries or cultural heritage organizations on their digital initiatives; 15 (38.8%, the majority of organizations with smaller staff sizes) do not collaborate, and two organizations did not know.

There were 36 total collaborative, or partner, institutions listed. The digital collaboratives that respondents organize or participate in, or institutions that they collaborate with, include: *Publication of Archival Library & Museum Materials (PALMM)* with 11 collaborators; *Central Florida Memory* with 2 participants; and the *Digital Library of the Caribbean* with 4 participants. Partner institutions most often named were the University of Florida and Tampa Bay History Center. The 22 institutions who collaborate reported participating in a wide variety of activities, including:

- Promotion of Shared Collections: 16 (72.7%)
- Using common sets of standards and best practices for metadata: 15 (68.2%)
- Using common set of standards and best practices for content creation: 14 (63.6%)
- Submitting collaborative grant proposals: 12 (54.5%)

April 18, 2014

- Shared digital asset management systems: 11 (50%)
- Shared digital preservation programs: 10 (45.5%)
- Creating exhibits from shared collections: 9 (40.9%)
- Shared institutional repositories: 5 (22.7%)

Those who collaborate were asked to rank the importance of a number of collaborative digitization project goals. All of the goals listed received a majority of ranking at the, "most important" level. By number of votes, these were the most popular goals:

- To increase visibility and expand the audience for our collections and organization: 14 responses
- To share collection resources among members of a collaborative, providing visitors and end users access to collections: 12
- To identify and share standards and best practices for improving access to collections: 11
- To share the cost of developing digital infrastructure and collections: 11
- To participate in a grant that supports collaborative initiatives: 10
- To develop the capacity and resources to digitize materials: 10
- To share technical resources including server space; infrastructure: 10
- To provide implementation assistance, ongoing training and consulting on digitization: 9
- To identify and share standards and best practices for the digitization of different types of media: 9
- To expand personal and organizational networking capabilities: 6

Digital Preservation

Thirty-nine institutions responded to questions in this section of the survey. 20 responding organizations (51.3%, including twelve academic libraries) have a digital preservation program; 15 (38.5%) do not, and 4 institutions do not know. Thirty-six organizations (92.3%) expect to retain digital collections for the long term (more than 10 years).

Institutions are funding, or intend to fund, preservation of digital collections through a line in their operating budgets: 16 (41%); through their IT budgets: 12 (30.8%); through grants: 11 (28.3%); 8 don't know; 3 don't plan to fund a digital preservation programs; and 2 plan to fund it through their preservation budgets. Academic libraries are the most diverse in their funding sources for digital preservation. "Other" answers to this query were comments such as: "leveraging shared resources and collaboration with other groups," and, "through the *Florida Digital Archives*."

According to *Trustworthy Repository Audit and Certification: Criteria and Checklist*; backup is the, "periodic capture of information to guard against system or component failure, or against accidental or deliberate corruption of the system." When asked if they backed up their collections, 37 institutions (94.9%) said yes, and 2 did not know.

Most organizations, 16 (43.2%) perform daily backups; 8 (21.6%) backup weekly; 3 (8.1%) backup quarterly; and 2 (5.4%) backup monthly. However, 1 institution said they backed up once, never backed up or did not know. Other answers showed that backup schedules varied by collection systems.

April 18, 2014

Thirty-seven of the organizations that backed up their files most often stored them in in-house systems their institutions manage: 31 (83.8%); participate in a shared repository system that is responsible for system backups: 10 (27%); outsource backup to private contractors: 8 (21.6%); or don't send their backup files offsite: 1 organization. Of the 9 institutions that responded "Other," 3 each said they used hard drives stored offsite, or at multiple sites.
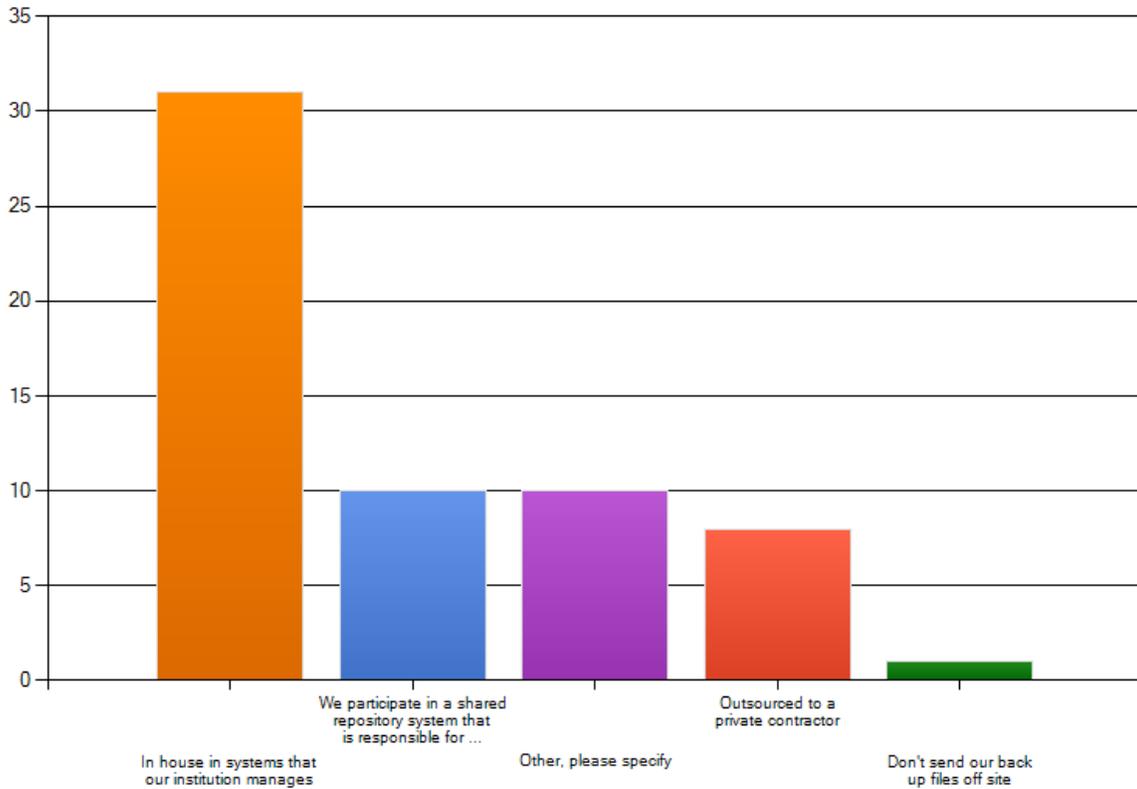
Digital preservation strategies implemented by the responding institutions include:

- Data backup: 32 (86.5%)
- Institution participates in a shared digital preservation repository: 8 (21.6%)
- Replication: 8 (21.6%)
- Migration: 8 (21.6%)
- Refreshing data: 5 (13.5%)
- Maintenance of legacy equipment: 5 (13.5%)
- Institution <u>subscribes</u> to a digital preservation service (*Portico, OCLC Digital Archive*): 3 (8.1%)
- Institution <u>operates</u> a digital preservation repository: 3 (8.1%)
- None: 2 (5.4%)
- Emulation: 1
- Don't know: 1

Media used for storage in digital preservation procedures include:

- Online magnetic media (servers, networked hard-drives): 30 respondents (81.1%)
- Optical media (CD, DVD): 15 (40.5%)
- Removable magnetic media (disks, zip disks): 13 (35.1%)
- Tape: 9 (24.3%)
- Don't know: 3 (8.1%)

April 18, 2014

If you back up your files, where are they stored? Select all that apply.

Respondents were asked what preservation enabling or digital preservation service or software they participate in or have implemented, or that they are investigating. 8 organizations, all academic libraries, participate in the *Florida Digital Archive*s; 6 in *Islandora*; 3 in the *OCLC Digital Archive*; 2 each in *Hathi Trust, LOCKSS/CLOCKSS*, or *Portico*; and 1 each in *D-Space*, *Fedora* or a *LOCKSS* Private Network. For services being investigated, 5 organizations are looking at *Islandora*; 4 each at *Hathi Trust* and *DuraSpace's DuraCloud*; 3 each at *Florida Digital Archives (DAITSS), OCLC Digital Archive* and *Fedora*; 2 at *Archivemetica* or *D-Space*; and 1 each at *Chronopolis* or *Tessella's Preservica*. Those organizations with larger budget sizes are investigating the largest number of these digital preservation solutions.

Finally, the survey asked responding institutions to share other comments or needs regarding their institution's digital projects or activities. 12 organizations answered with comments about the lack of funding for digital programs; concerns about costs of specific services; the need for a, "statewide historical cloud;" and 3 additional comments on institution-specific collections.

April 18, 2014

Summary:

Top administrators and professionals at a wide variety of cultural heritage institutions participated in the *Florida Statewide Digitization Survey* in early 2014. Digitization activities are underway at institutions with a wide range of staff and budget sizes.

Many of these cultural organizations have been creating digital collections for over a decade. However, the majority do not have documented policies addressing digital collections in the areas of: mission and goals; collection development; emergency preparedness; exhibits; digital preservation; strategic planning; or rights and licensing. Additionally, the staff sizes and operating budgets for the digital collections initiatives in Florida institutions are relatively small.

Access, study and use of materials are the main reasons institutions are creating, and/or acquiring, digital collections. Organizations are selecting materials for digitization because of strong local interest, fragile or deteriorating condition, high value and heavy use.

Photographic prints and flat works on paper are the materials most often used as sources for creating digital collections, and a vast majority of these collections are created in-house, rather than outsourcing digitization to a vendor. In addition, in almost every subject area related to Florida, photographs are the top material format to have already been digitized. Top born-digital formats created or acquired included: documents; photography and other still images; digital video; and digital audio.

A concern among the findings is that the majority of institutions answering questions about digital content creation standards are not using well established national standards, but are instead utilizing in-house, local and vendor recommended standards.

Few Florida organizations currently have large digital collections. Out of those responding to the survey, only six organizations hold over 10 terabytes of material. However, almost all institution types and sizes anticipate at least a modest gain in collections size between 2014 and 2017.

When asked about collections that could be digitized in the future, respondents expressed interest in digitizing all formats of material, including: photographs; maps; manuscripts; audio; videos; books; and texts, which are currently the most popular formats, especially if the materials focus on the themes of local and Florida history.

Surprisingly, 34 institutions have digital asset management systems, while 33 do not. The adoption rate in Florida is lower than in many states and regions the consultants had previously studied. Institutions are utilizing national standards, controlled vocabularies and schema when creating metadata for their digital collections; they are also utilizing recognized best practices in many aspects of copyright and rights management. One area that could use improvement is updating institutional Deeds of Gift/Donor Agreements, to include digital content creation and digital preservation activities.

Many Florida institutions are collaborating with other cultural heritage organizations on digital projects and initiatives; promoting shared collections; and using common sets of standards and best practices for

April 18, 2014

metadata and content creation. The main goal of this collaborative work is to increase visibility, as well as expand the audience for the organization and their digital collections.

Just over half of the organizations responding to questions on digital preservation have developed a program in this area, which is a very positive finding. Backup practices are widespread and regularly implemented. However, one major concern about digital preservation is that a majority of organizations store their backup files in-house, in systems their institutions manage. This is a concern because geographic distribution of digital collections for preservation is an established national standard. In Florida, which often experiences natural disasters, the idea of storing copies of digital collections in a geographically distant location may help in disaster recovery and business resumption. Another concern is that use of optical media carriers, such as CDs and DVDs, is still relatively widespread for digital preservation storage, although the life expectancy figures for these types of media are constantly being reduced as new reports of testing are published.

Recommendations:

Florida's cultural heritage organizations are making available to Florida's residents a significant body of digital content on a wide range of topics. Several challenges exist moving forward.

- Expanding participation: Expanding the range of cultural heritage organizations that are engaged in creating and making their digital collections available to Floridians should be a key focus of any future efforts. Bringing in more cultural heritage organizations will create a wider depth and breadth of content, as well as diversify the awareness and use of the collections.
- Public access to digital collections: The methods by which organizations can provide their collections to the public must be expanded. Developing a single point of access to Florida's digital collections will help meet this goal.
- Digital Asset Management Systems: Just 50% of the organizations creating digital content have digital asset management systems that support access to their collections. Developing solutions that will expand the use of digital asset management systems will allow organizations to improve the local management to collections, as well as support the collaborative sharing of collections, both nationally and internationally.
- Training: The survey identified a lack of a programmatic approach (which could be implemented on a regional or statewide basis, or by type of organization, such as museum/archival/library) that support small and medium sized organizations in their efforts to digitize their collections, or build digital collections through acquisition. Programmatic elements include: training/education in digital standards and best practices; platforms for access to digital content; digital preservation; and promotion of digital collections.

In addition to the above recommendations, the following recommendations relate to both individual organizations, as well as possible collaborative digital efforts, within Florida's cultural heritage community:

April 18, 2014

- Training: To expand the participation of additional cultural heritage organizations, as well as the use of national standards and best practices, broad based training is required. The survey identified a lack of a programmatic approach to support organizations in creating or acquiring digital collections.
- Moving from project to program: Over the past two decades, Florida's cultural heritage organizations have been creating digital collections. However, few policies have been developed to support these collections. Across the state, standards and best practices for the creation of digital content are diverse, from use of national standards, to vendor provided specifications, to a complete lack of standards. To meet the needs of: researchers and scholars; K-12 students; distance learners; as well as local historians, quality digital collections and associated metadata must be created and broadly implemented.

Programmatic elements to be developed include: training and education in standards and best practices; platforms for access to digital content; digital preservation; and promotion of digital collections.

Statewide guidelines and best practices in many areas of digital practice need to be created. Providing wide access to existing policy models from Florida cultural institutions can assist in the development of new statewide guidelines. Additionally, utilizing existing sample policies and documents that deal with updating institutional Deeds of Gift/Donor Agreements to include digitization, and digital preservation, as something that could be done with gifts and donated materials, could be beneficial to many Florida institutions.

Subjects for future digitization: Based on the survey, a focus on building statewide collections emphasizing subject areas, such as local and Florida History, is a popular direction among respondents. These subjects can be further defined, or loosely defined, as digital initiatives move forward, but these topics are certainly areas of interest.

APPENDIX A

POTENTIAL SUBJECTS WHERE FLORIDA COLLECTIONS COULD BE DIGITIZED

- Florida History
- Beaches
- Fishing
- Boating
- Geology
- Natural Resources
- Environment/Ecosystems
- Space Exploration
- Native Americans
- Civil Rights
- Winter Visitors
- Energy Resources
- Water Resources
- Tourism
- Entertainment Industry
- Agriculture and Ranching
- Art and Architecture
- Music
- Literature
- Climate
- Business
- Local History
- Religion
- Ethnicity
- Politics and Government
- Sports
- Transportation
- Technology