Museum Technology Landscape 2018: Discovery and Findings

Original research and report produced by
Nik Honeysett and Julia Falkowski - January 2018

Adapted by Laurie Gemmill Arp and Megan Forbes

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report summarizes a study of collections management systems and practices within the museum community conducted by LYRASIS with funding in part from The Andrew W. Mellon Foundation. The study was comprised of surveys and individual interviews. This report has been adapted from the original report to remove proprietary, confidential and respondent-identifying information. Opinions expressed in the report are those of the authors and collaborators, and do not necessarily represent those of The Andrew W. Mellon Foundation.

The study covers information about the status of collections, collections management practices, digitization, software selection processes, technology support, and collections management strategy, specifically within the museum field.

INTRODUCTION

In November 2017, LYRASIS contracted with the Balboa Park Online Collaborative (BPOC) to gain understanding of the collections management world. This included looking at the technology use and practices currently playing out in the museum community.

To pursue these questions, from November 2017 to January 2018, BPOC employed surveys and interviews to solicit feedback from the collections management community at large. While the survey was distributed internationally, respondents are primarily from the U.S. and include curators, librarians, and developers, but most commonly collections managers. The body of this report provides a summary of the surveys and interviews. The appendices contain the survey questions. Datasets of the original survey responses are available on the LYRASIS website at: http://www.lyrasis.org/Leadership/Pages/Museum-Collections-Management.aspx. Note that proprietary, confidential and respondent-identifying information has been removed from these datasets. As part of the same overall project, Ithaka S+R conducted several case studies exploring non-textual collections management on university campuses. High level results from that analysis are also included in this report.

This engagement was led by BPOC's CEO, Nik Honeysett, supported by Julia Falkowski, Project and Production Specialist. The core project team, which met each week to review progress, included Roger Schonfeld, Director of Library and Scholarly Communication Programs at Ithaka S+R, who focused on evaluating academic audiences, and a number of LYRASIS and CollectionSpace representatives, most notably, Director of Collection Services and Community Supported Software Laurie Arp and CollectionSpace Program Manager Megan Forbes.

The survey was organized into seven sections focusing on an institution's Collection Management System (CMS), but also pertinent metrics concerning practices and collections-related data: You and Your Institution, Your Collection, Your CMS Selection Process, Your CMS Implementation, CMS Support, Collections Management Strategy, and Final Thoughts.

The main section of this report is the *Survey Findings and Analysis,* which provides a detailed interpretation of the survey data.

METHODOLOGY

The study was conducted to gain a broad view of collections management and collections management systems in the museum and academic fields. The core data gathering instrument was a general collections management survey, the findings from which were elaborated upon through interviews.

BPOC began the evaluation process by crafting the survey. With input from the project team, and trusted connections in the Balboa Park collections community, BPOC developed a set of questions that they believed would provide the most valuable results. To increase completion rate and usability of data, BPOC avoided open-response questions as much as possible, split the questions into digestible segments, and let users know they could enter their email to receive a link to complete the survey later. The survey tool employed, SurveyGizmo, estimated that the survey would take about 20 minutes, and was ranked green (low) on its *Fatigue* score and green (maximum) on its Accessibility score.

In December 2017, BPOC deployed the survey first to its collections management connections in Balboa Park for testing. After this first round, the survey was primarily distributed to communities in the United States, including Canada, but also to the U.K. and Australia. Requests went out on BPOC social media channels - Facebook, Twitter, and LinkedIn; the museum technology listserv MCN-L; the Museums and the Web online discussion board; and the U.K.-based Museum Computer Group. BPOC also reached out directly to collections management and registrar connections who might be interested.

The interview portion of the evaluation was informed by the survey respondents' answers. BPOC requested follow-up interviews with those who had both indicated their willingness and who wrote an answer in the *Final Thoughts* open-response question at the end. BPOC scheduled half-hour phone interviews with these individuals, to elaborate on specific survey answers.

The survey received 214 responses for 66 completed surveys, and 23 interviews. Survey respondents were largely based in the United States, but some international respondents provided valuable context as well.

The sample size for the survey implies that a single institutional entry will affect the results by approximately ±1.5%.

SURVEY FINDINGS

Following is a detailed reporting of the survey data.

RESPONDENTS
This word cloud represents titles of respondents, and while it does not make much sense to provide analysis or interpretation on these, it is interesting to note that curators and directors are featured. The word "amp" is a feature of the survey tool converting the symbol "&" to a word and its presence in the word cloud indicates a significant number of combined job titles, e.g. Collections & Database Manager.

INSTITUTIONS

Over half the responses are either an art museum or have an art collection, followed by history, and represent all sizes of museums. 9% are under $100K annual budget, one quarter are above $25M annual budget, and 6.5% have over 1,000 staff members. More than half of institutions report having an archives. One third represent small museums, defined as $300k budget or less and up to five staff. While an accurate comparison of how closely this survey represents the field, one data point is that, of the approximately thirty thousand museums in the U.S., twelve thousand are considered small, i.e. one third.

IT STAFF

Two thirds of respondents have 0 or 1 IT person. There appears to be a pervasive lack of appropriate IT support in terms of budget and resources across the museum field. A quarter of respondents have more than three IT staff.

BUDGET AND FINANCES

The average budget of respondents is almost $17M, with a median at less than $5M. CMS total cost of ownership estimates cover a wide range; a slight majority reports under $10,000 annually, but estimates extend to more than $50,000. The chart below shows the spread of the total cost of ownership, defined as annual costs associated with installation, licensing, subscription, hosting, support, etc., with a calculated average of $15,000. Respondents were also questioned on how they budgeted for their initial CMS setup and installation. While almost one half budgeted for it, one quarter relied on grants or gifts to help finance their migration and set-up. Like collection digitization, which is most often a grant-driven project, this results in a portion of the field operating with a somewhat opportunistic approach to technology improvement and upgrades, rather than operational funding, even for a core activity such as collections management.
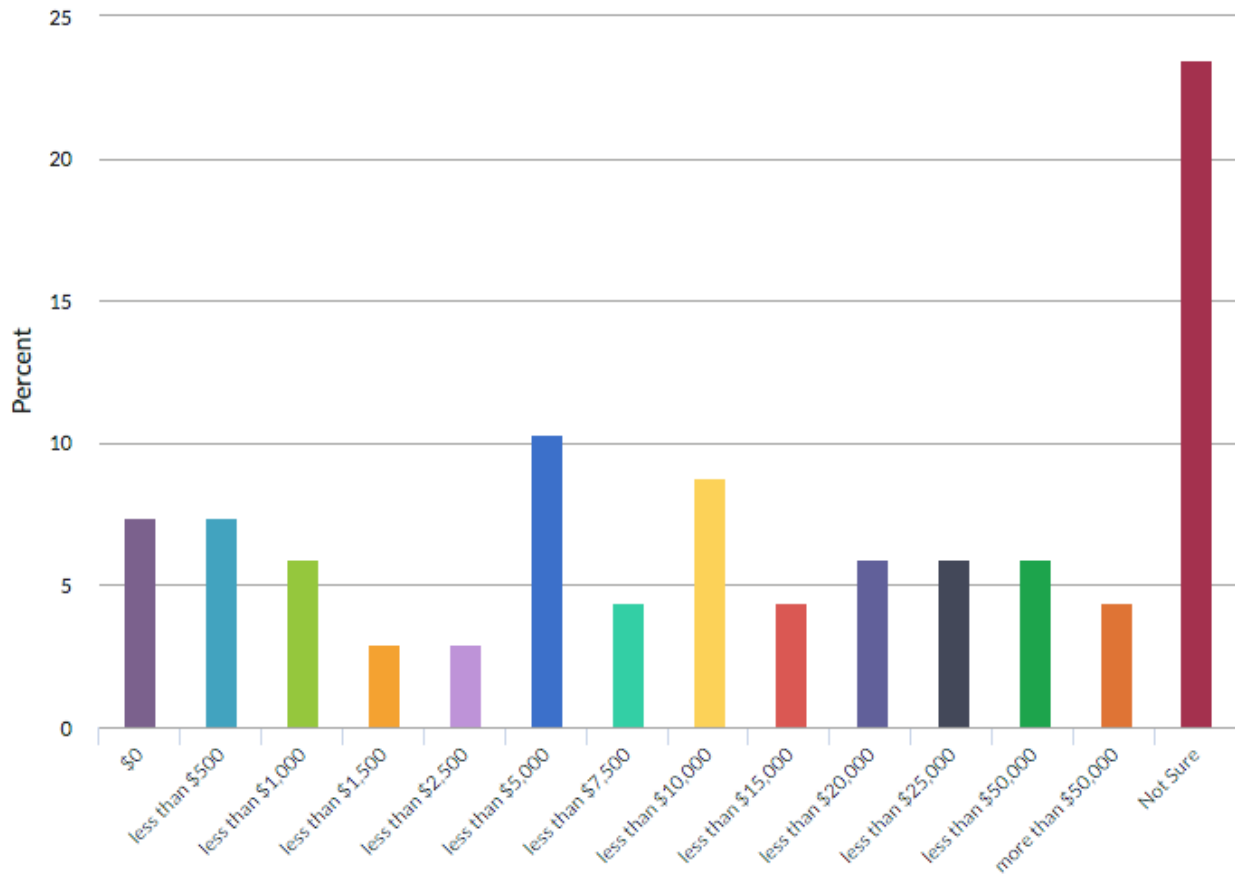
THE ANNUAL TCO OF CMS LICENSING/SUBSCRIPTION/HOSTING

COLLECTION SIZE

The survey represents a varied array of collection sizes, but it biased somewhat toward the low end with a slight majority being below 20,000 objects.

COLLECTION ELECTRONICALLY CATALOGUED

Approximately two thirds of survey respondents have two thirds or more of their collections catalogued electronically. Less than 25% have 100% of their collections cataloged electronically. Cataloging rates are below 1,000 per year for three quarters of respondents, and a majority of collections are stored onsite.
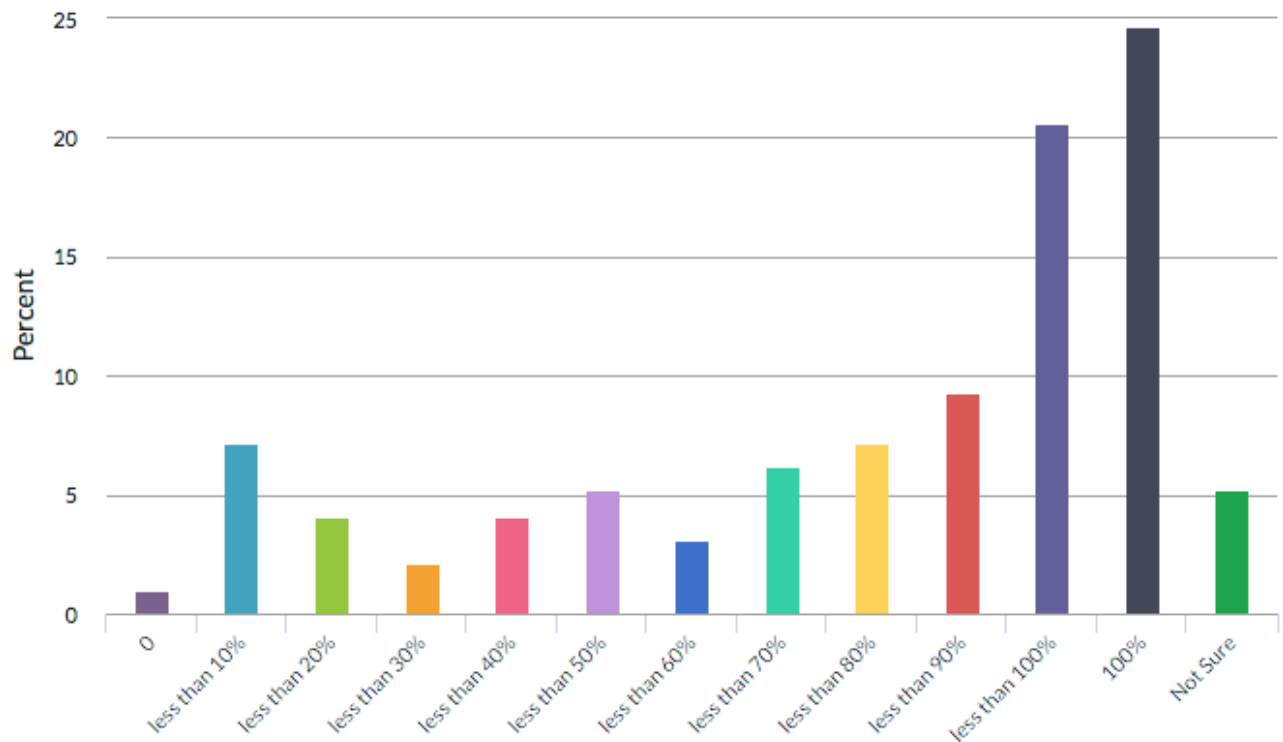
PERCENTAGE OF COLLECTION CATALOGUED ELECTRONICALLY

OBJECT ACQUISITION RATE

Almost two-thirds of respondents have an acquisition rate below 250 objects per year, a 90% majority have rates below 3,000 per year and several outlier respondents with rates above 10,000 per year.

COLLECTIONS ON VIEW AND ONLINE

The majority of respondents have less than 10% of their collection on view, and almost three quarters have less than 10% of their collections available online. The chart below characterizes how respondents are currently making their collections available online, which is largely by highlights through a custom portal.
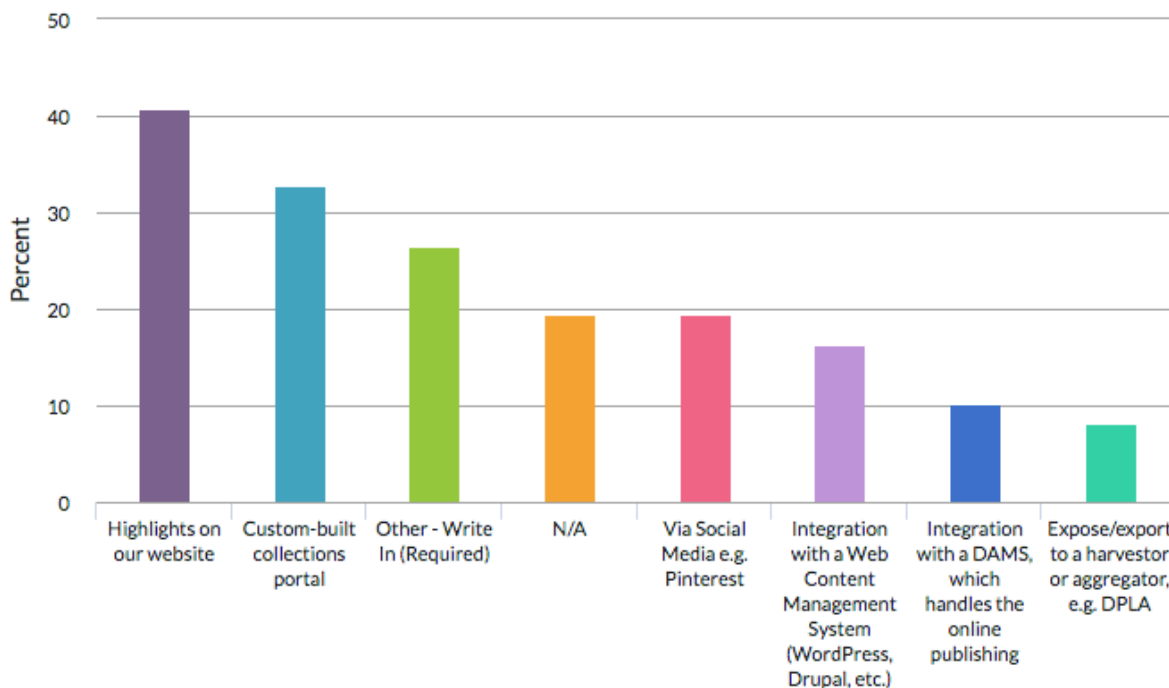
HOW COLLECTIONS ARE MADE AVAILABLE ONLINE

Strategies and timelines associated with delivering collections online present a mixed and broad view. A majority expect to deliver their collections on their website through a module within their CMS, i.e., rely on their vendor to provide this. This is not surprising given the availability of PastPerfect Online, eMuseum, and WebKiosk. Approximately one third are looking to deliver via a direct integration with their Web Content Management System, and/or rely on harvesters or aggregators to provide that access.
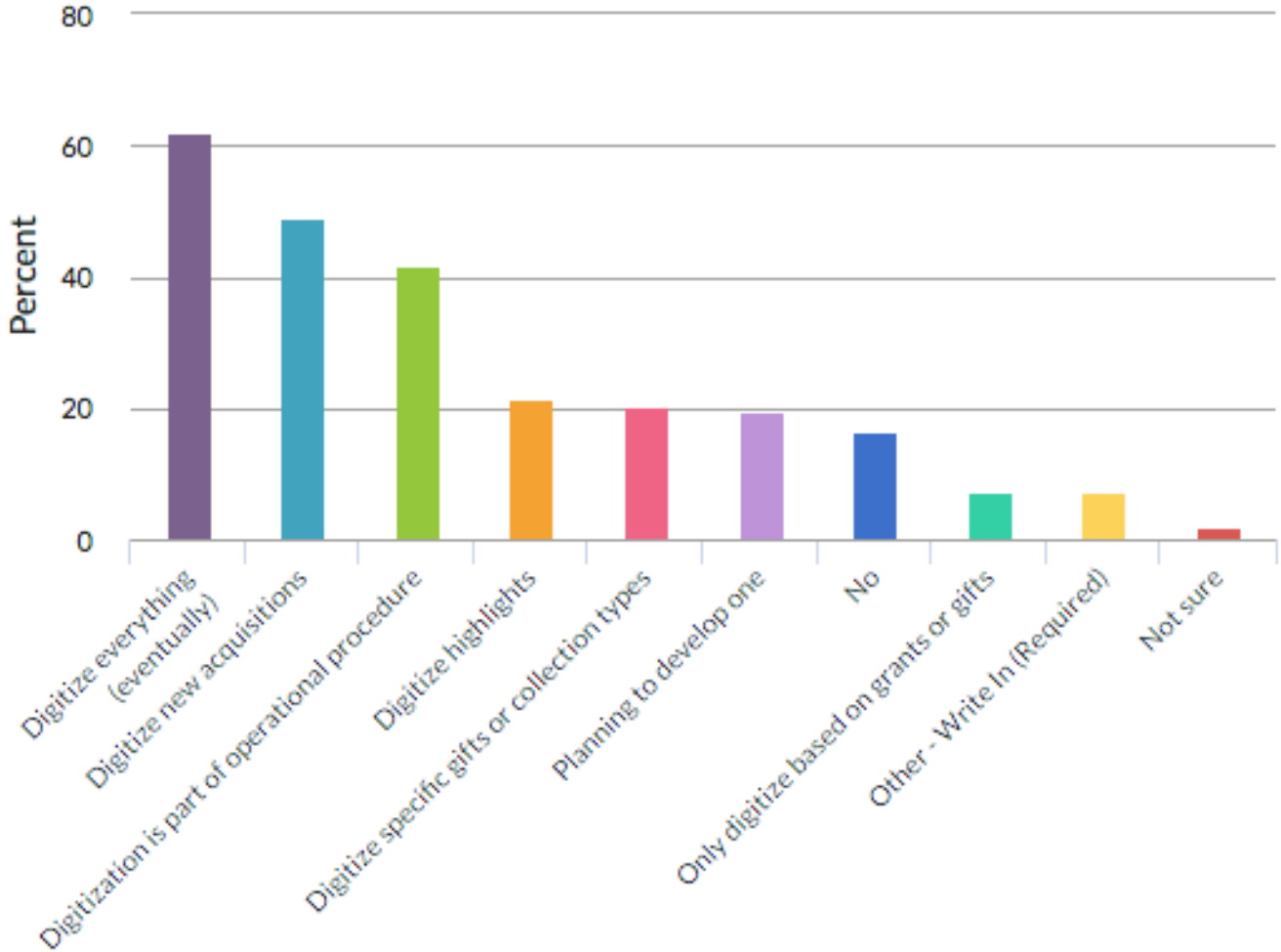
Institutional policy with respect to online access was also questioned. Two thirds described some form of controlled access approach, with 8% offering unrestricted access including commercial use. For controlled or protected use, one quarter protect through thumbnails or low resolution only, another quarter control use to high resolution images through education-only licensing, stated restrictions or watermarking, and one fifth state a variable form of access based on the highest resolution they can based on rights.

Approximately one third of respondents have not yet made any of their collections available online. The timelines expressed to do so describe long-term plans of five to ten years representing approximately 12% of all respondents, another 12% have no explicit plans or no faith that the institution will be able to deliver such a project. The remainder are looking to deliver in the short-term, within six to 24 months.

DIGITIZATION AND IMAGE MANAGEMENT

Approximately half of the represented collection objects have no analog (e.g., slide or photograph) or digital surrogate. With the vast majority of collections not on view, there is no way for the world to see what an object looks like without seeing it in storage. With institutions only slightly exceeding their acquisition rate over their digitization rate, there is a comprehensive mismatch between stated digitization goals ("to digitize everything at some point") and an institution's ability to deliver on them.

DIGITIZATION STRATEGIC GOALS

On average, institutions increase their collection size by a single percentage point per year, their cataloguing and digitization rates are both about half a percentage point higher. Given the extent to which a significant number of collection objects remain undigitized, it will take decades if not centuries to comprehensively represent collections in the digital environment. The likelihood is that this is an unattainable goal.

The chart below captures the state of photographed (analog) and digitized collection objects for 61 respondents, i.e. each stacked bar describes the percentages of no analog or digital representation for a collection object, photography, high and low-resolution capture. The chart is ordered by no analog or digital representation, meaning the institution on the far left has high resolution capture for its entire collection, while the institution on the far right has only a tiny percentage of photography or digital capture.

Page 9 of 18

DIGITIZATION STRATEGIC GOALS

On average, institutions increase their collection size by a single percentage point per year, their cataloguing and digitization rates are both about half a percentage point higher. Given the extent to which a significant number of collection objects remain undigitized, it will take decades if not centuries to comprehensively represent collections in the digital environment. The likelihood is that this is an unattainable goal.

The chart below captures the state of photographed (analog) and digitized collection objects for 61 respondents, i.e. each stacked bar describes the percentages of no analog or digital representation for a collection object, photography, high and low-resolution capture. The chart is ordered by no analog or digital representation, meaning the institution on the far left has high resolution capture for its entire collection, while the institution on the far right has only a tiny percentage of photography or digital capture.

COMPARISON OF PHOTOGRAPHED AND DIGITIZED COLLECTION OBJECTS

Approximately half of respondents do not have a managed environment for their digital assets; the remainder is split between a dedicated digital asset management system (DAMS) and the management capabilities within their CMS to manage digital images.

COLLECTIONS MANAGEMENT STRATEGY

The chart below details responses to whether an institution has strategic goals in place for collections management, almost one third of respondents describe having no roadmap or currently in the planning stages. An overwhelming majority has collections management goals in place of varying degrees of detail and support. Just under half cited broad institutional goals, just over half cited specific goals, but only one third had any detail in terms of deadlines, budget and resources to support those goals.
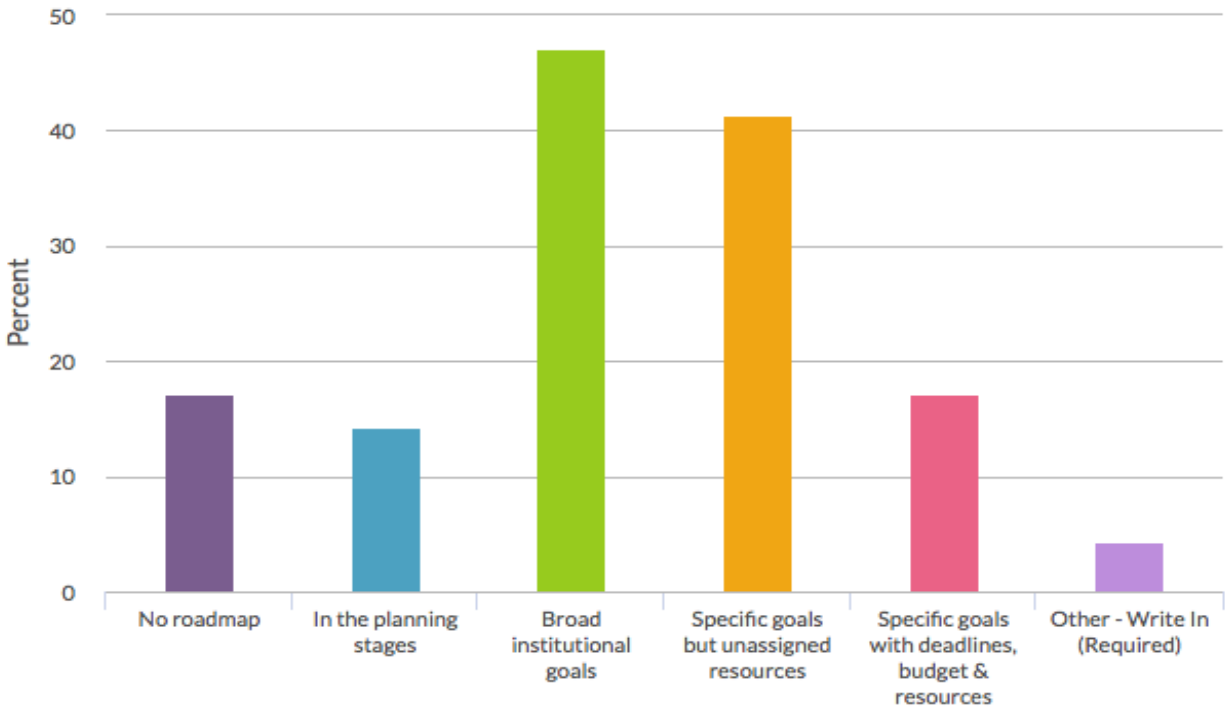
A majority of respondents agreed (though not strongly) that their CMS: is meeting their collections management needs; is easy to administer; is easy to support; and is easy to upgrade, though there were significant numbers who rated these factors in the "disagree" and "neither agree nor disagree" categories. The majority of system integration needs fall in the "disagree" and "neither agree nor disagree" categories. Ratings for the ease of learning and use fall into the "disagree" and "neither agree nor disagree" categories.

A majority of respondents seek support from their vendors infrequently, mainly looking for "how to use", "how to administer" and technical installation help. Vendor support is on-the-whole meeting expectations but a majority said they neither agree nor disagree that their support is better than for other CMS products they have used. A majority tend to rely on their vendor rather than peer or communities for support, agreed (though not strongly) that their vendor is responsive to feature requests and bug fixes, and neither agreed nor disagreed that their vendor clearly communicates their product's roadmap.

CMS SELECTION PROCESS

The CMS selection process generally involves collections staff but is not always led by them; leadership and IT are involved to a lesser extent. Leadership is the primary final arbiter of the decision, followed by collection staff, indicating that price, possibly at the expense of addressing required functionality, is a strong determinant. The chart below identifies that half of staff involved in CMS selection no longer work at the institution, which under normal circumstances would not be a cause for concern, however when mapped to the reported lack of formality around product selection, where peer recommendation, prior familiarity or "other informal methods" equate to over three quarters of institutional "process", frustrations with the incumbent system are to be expected. Only 11% of respondents claim to have followed a formal process to select their CMS.
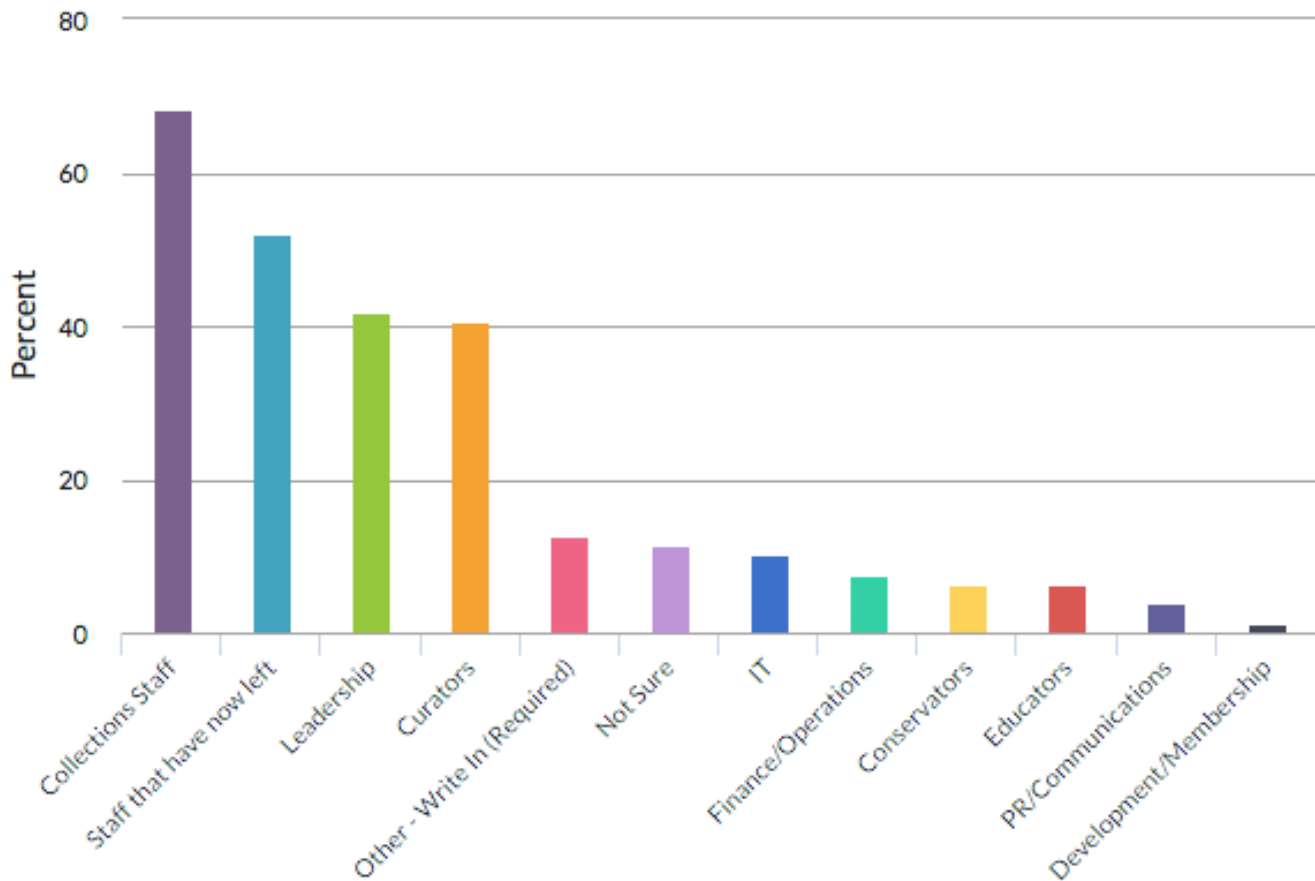
STAFF INVOLVEMENT IN THE CMS SELECTION PROCESS

The survey asked respondents to rate the importance of a wide range of criteria in making their product selection, including the availability of an on-premises server, vendor and community support, the user interface, public access, and media management. Average rankings of criteria show querying, media management, metadata management, and object management to be the top concerns for most respondents. The least important criteria were elements such as rights and reproductions management, a formal RFP process, and the requirement that the CMS be a desktop application.

During the selection process respondents tended to agree (though not strongly) that it was easy to find out information about vendors and their products; that implementation generally took more time and effort than they had anticipated; and a majority strongly agreed that CMS implementation requires more planning, preparation and training than anticipated. Post implementation, most respondents agreed (though not strongly) that the process they used resulted in the best solution and a majority agreed (though not strongly) that the CMS functions in the way it was presented by the vendor.

CMS INSTALLATION

The majority of respondents have been managing their collections electronically for about 15 years or more, and the majority have had their current CMS for about the same length of time. The chart below shows the spread of respondents' currently installed CMS products, which somewhat reflects the ranking of the CMSs installed

across the entire museum community. PastPerfect is recognized as having the largest installed base of a CMS at above 10,000[1], TMS about 1,600 and EmbARK about 400. TMS and EmbARK are used predominantly by art museums; their strong showing in the survey results confirms the museum type results where art is the largest representative museum. The results beg the question, why are art museums more inclined to respond to surveys?
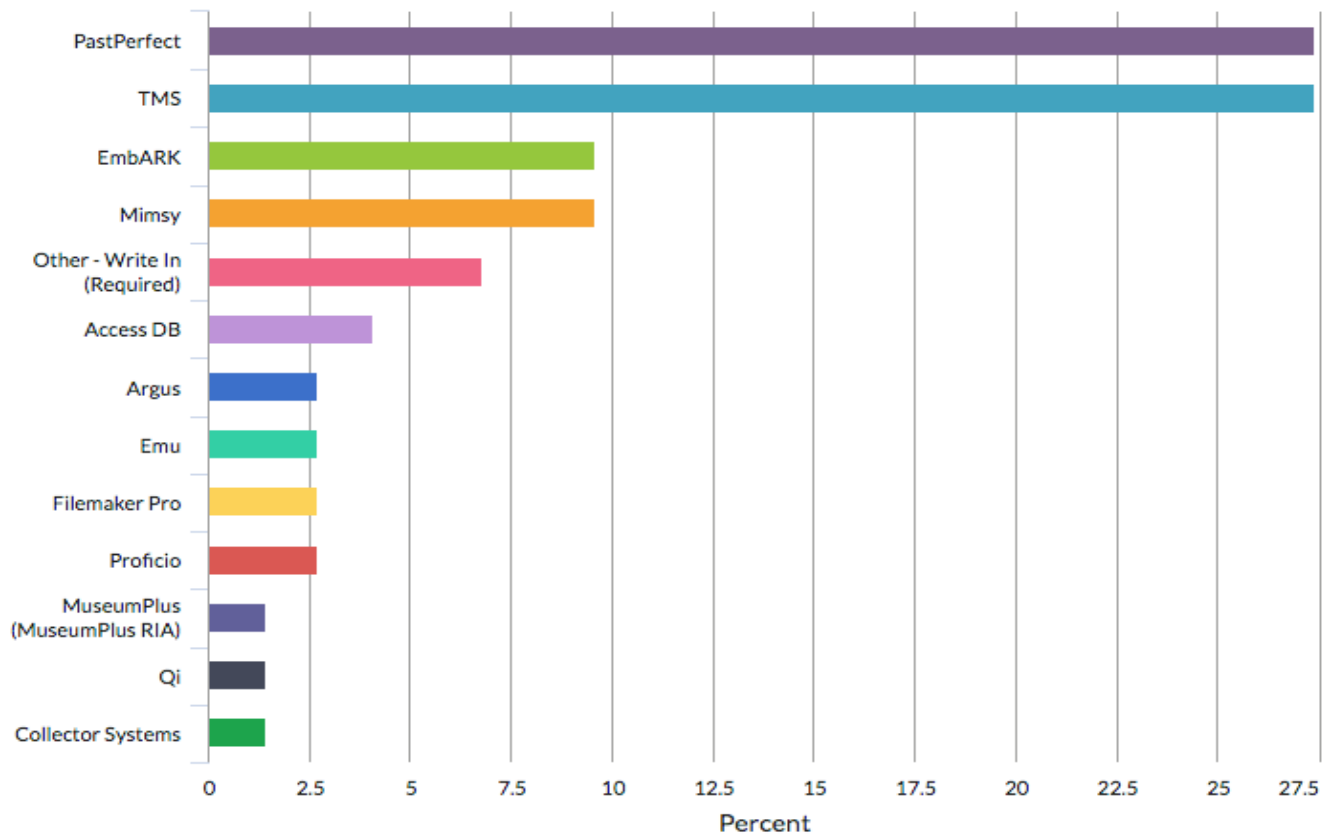
DISTRIBUTION OF CMS PRODUCTS

Two thirds of CMSs are installed on an institution's network. This follows on from the most popular CMS products, which are all networked-installed products from PastPerfect, TMS and EmbARK. One third of installations are installed and configured by in-house staff, the remainder by the vendor or third-party providers. One quarter have their CMSs maintained and supported by in-house staff, the remainder by the vendor or third-party providers. Overwhelmingly, staff are using the product multiple times daily or daily, with half of institutions having more than one dozen accounts.

Only a small percentage have integrations with other systems. These are primarily a web content management system (WCMS) or DAMS and a few preservation systems, although this was not well defined. A few respondents have manual or scripted integrations with membership and development systems. WCMS integrations tend to be scripted (some of the most-used products have collections-online modules, e.g., PastPerfect Online, eMuseum and WebKiosk), DAMS integrations are equally scripted or manual.

---

[1] http://www.museumsoftware.com - retrieved 01/24/2018

Encouragingly, the use of controlled vocabularies features significantly in the collections management usage, likely because many vendors support them, although during interviews respondents would prefer an integration-based support rather than loaded into their CMS. As one might assume, the Getty's offerings dominate significantly, with the Nomenclature for Museum Cataloguing and Categories for the Description of Works of Art (CDWA) showing healthy implementation.

Reporting capability is a "very important" functional requirement across the board, and the reports produced drive or inform significant activity across institutions, with loan show information and photography being two of the most significant. During the interview process, there was a fair amount of frustration expressed about the complexity and usability of reporting functionality.

Finally, less than ten percent of respondents strongly agreed in each of these categories: that their institution provides sufficient budget and resources to deliver on their collections management goals (a majority disagreed); that their institution has a clearly articulated collections management strategy (a majority agreed); that their CMS meets current collections management needs (a majority agreed); and that their CMS will support their collections management strategy (a majority agreed). However, almost half strongly agreed that their CMS is critical to their institution's collections management strategy.

INTERVIEW HIGHLIGHTS

BPOC completed 23 total half-hour interviews with survey respondents. In these interviews, BPOC asked respondents to clarify and elaborate on their survey answers.

PRODUCT FIELD INTERVIEWS

There was general dissatisfaction with CMS options and functionality. Organizations are particularly skeptical of proprietary programs that might consume their data, or charge for its extraction. Given the ability to easily access data, interviewees explained, they would build connections with other programs that would make a whole host of vital tasks and strategic goals more attainable; these include condition reporting, online publishing, and sharing collections with aggregators. One interviewee described this well by saying:

> "A CMS can't be everything, but it needs to play well with programs that need that information."

Currently institutions that are pushing the integration limits with their CMS are achieving it by creating intermediate data layers.

Many in the field are looking for broader functionality than what is currently offered. On the truly innovative side, this includes incorporating machine-learning into object tagging; incorporating museum collections with libraries and archives; and integrating an approval process that allows volunteers to catalog, and curators to approve, before any data becomes final.

ENVIRONMENT FIELD INTERVIEWS

In general, whether organizations relied on in-house, vendor, or third-party support, interviewees generally thought they were adequately supported technically.

Financial concerns remain a major factor throughout the field. While many believed their museum provides adequate resources to achieve formal and informal collections goals, many described leadership that wanted to support collections but was unsure how to do that or uninformed about what that might mean. A few people mentioned the value of tools that could help inform leadership of their work. While most interviewees were excited about digitization and providing access, others thought their resources and time would be better spent taking care of the physical collections objects, and that concerns of access should fall under the ownership of another, currently non-existent, role in the museum.

Multiple people mentioned the desire for a community that includes museums and vendors with a degree of commitment towards standardization, at least in how data is modeled. This would allow for greater integration and share ability, increasing access and helping organizations meet their broad strategic, forward-thinking goals.

Ithaka S+R also conducted site visits to the museums of three universities University at Buffalo (SUNY), University of Delaware and Yale University. These provided real breath in terms of museum size/scope, organizational structure, and university type. Broadly, we found that these museums were in very different places in terms of their collections management approaches, capacities, and requirements. While some cases suggested needs for extremely sophisticated new platforms beyond what is currently available, others had more of a need for low-overhead solutions that would allow more basic processes. In all cases, there was a need for collections management approaches that could better integrate museum collections and activities into broader collecting and curricular work at the parent universities.

LYRASIS

APPENDIX A – SELECTED SURVEY QUESTIONS

YOU AND YOUR INSTITUTION
- Your Title (only shared in aggregate)
- What type of Museum do you work in? (check all that apply)
- Approximately how many staff work in your museum (Full-time and Part-time)?
- Approximately how many dedicated IT staff work in your museum (Full-time and Part-time)?
- Approximately, what is your total institutional budget?

YOUR COLLECTION
- Approximately, how many objects are in your collection? (We are not concerned about your archive or library.)
- Approximately, what percentage of your collection is catalogued electronically?
- How many objects do you acquire per year on average?
- Approximately, how many objects are you cataloging per year?
- Approximately, what percentage of your collection is located at your primary site(s)?
- Approximately, what percentage of your collection is on view in your galleries on average?
- Approximately, what percentages of your collection have no surrogate image, or have been photographed or digitized for access? (total should equal 100%)
- Approximately, how many objects are you digitizing per year?
- Are you digitizing by shooting/re-shooting the original object or by scanning existing photographs, transparencies or other analog media?
- Do you have a digitization strategy (check all that apply)?
- Do you use a separate application or cloud service to manage your digital images?
- What percentage of your collection is available online?
- How have you made your collection available online? (Check all that apply)
- Do you have plans to make your collection available online?
- How will you make your collection available online? (Check all that apply)
- How many people are directly involved in collections management and support?

YOUR CMS SELECTION PROCESS
- Do you have a strategic roadmap for collections management, i.e. cataloguing, digitization, etc.? (check all that apply)
- Which staff were involved in the selection of your CMS? (check all that apply)
- Who led the selection process? (check all that apply)
- What process was followed to select your CMS? (check all that apply)
- Who had final say in your selection?

YOUR CMS IMPLEMENTATION
- How long have you had your current CMS?
- If this is not your first CMS, what CMS did you migrate from?
- Who provides technical support for your CMS installation?
- How often do you or your staff log into your CMS?
- What percentage of your collection is managed in your CMS?

- Does your CMS currently integrate or connect with any other systems? By what mechanism?
- What type of staff use your CMS? (check all that apply)
- What percentage of records have an associated media object, e.g. image, pdf, spreadsheet, etc.?
- What percentage of records have associated metadata that promotes public access or discovery?
- What percentage of records have interpretive or educational copy for public access?
- What types of media do you associate with your records? (check all that apply)
- What controlled vocabularies do you use to support your use of your CMS? (check all that apply)

APPENDIX B - FIELD QUOTES

The following are a selection of quotes from the surveys that highlight pertinent or more widely held thoughts and comments:

"As users, we bring our broader tech expectations to our CMS (it should work like Google, like my iPhone,...). The flexibility of systems to expand and meet evolving user needs continues to be paramount, e.g. by becoming device agnostic, browser-based, increasingly configurable, etc. The better our vendors can keep pace with the overall dynamic tech environment, the better for us."

"I am increasingly skeptical the dollars thrown at digital content support the care of the physical objects in our collections. We have very poor support for digital infrastructure, however, if I was given a pot of money, I can identify a number of physical collections needs that I would first put money toward. However, I seem to see more spending on digital content, and greater effort toward that and less attention to the basic needs of collections."

"Sustainable, i.e., more widely used, open source museum CMS solutions are sorely needed to compete with the veritable monopoly that certain enterprise software companies have held for too long. The irony of open source software is that most institutions can't afford it, in terms of being able to fund the technical staff/consultants needed to launch, maintain and integrate it with other platforms (web, DAMS, etc.)."

"All of the CMS I am familiar with are still designed and working on data structures that were designed before the advent of the internet, before big data, before AI, before machine learning. Advances in the past 15 years require a complete rethinking of what collection information should be collected, how this done, how it is shared, and how it is stored. In our institution we have moved recognized that our collection data needs to move beyond the bounds of the standard relational (i.e. SQL) based CMS. We need to move to graph focused data structures as it more readily reflects reality of the data and its use."