Heritage Health Information 2014 Survey

Section B. Collections of Information Employing Statistical Methods

B.1.   Universe, Sample Design, and Response Rate Estimation

The target universe for the Heritage Health Information 2014 Survey is all physical institutions in the United States with collectable holdings (i.e., museums, libraries, archives, historical sites/societies, and archaeological repositories or scientific research organizations). All institutions holding medium to large size collections and all archives and archaeological repositories/science research organizations with small size collections will be identified and included in the study. All libraries, historical societies, and museums with small size collections will be identified, but, due to the large number of these institutions, samples within each type will be randomly selected for inclusion in the study. Our goal is to obtain an overall response rate of 28%, however it is anticipated that response rates will vary depending on the sample group (70% for institutions with large size collections; 50% for mid-size collections; and 40% for institutions with small collections).

B.1.1. Universe and Sample Design

The universe for the survey includes all physical institutions with the responsibility of holding non-living collections in the public trust, such as books, photographs, moving images, recorded sound, digital materials, art and historic objects, archaeological collections, and natural science specimens. The institutions include libraries, museums, historical societies/sites, archives, archaeological repositories and scientific research organizations. This study does not include historic structures or living heritage, such as performing arts, or living collections in institutions such as zoos, aquariums, and botanical gardens. Other exclusions from the study universe include elementary and secondary school and two-year college libraries, branch public libraries, hospital libraries, and prison libraries since they do not hold rare or special collections. In addition, county clerk offices, law firms, newspaper, corporate, engineering firm libraries, and for-profit organizations are not included.

Organizational entities operating under a parent institution will be accounted for by the parent institution only. For example, a museum with a library will complete the survey for its museum and library collections. Systems of collecting institutions that have central collections control and preservation practices, such as a library system within a university, will be asked to complete the survey for the main library and departmental libraries. However, professional schools, such as a university's business school, medical school, or law school or university museums and departmental collections (archaeology or sciences) are often not centralized in administration, so they will be included in the universe individually.

Background. In 2001, the Institute of Museum and Library Services (IMLS) partnered with Heritage Preservation to develop and conduct the Heritage Health Index 2004 study (HHI 2004). The first task in developing the study was conducting a literature review of ongoing and previous studies on institutional preservation. It reinforced that no studies had addressed the breadth of U.S. collecting institutions and all the materials they are responsible for preserving. A

bibliography of consulted surveys and relevant publications is listed in Appendix A. Heritage Preservation established an Institutional Advisory Committee of 35 professional associations and federal agencies representing collecting institutions and working groups comprised of 66 members to advise on the survey development and implementation of the HHI 2004. Advisory Committee organizations and types of working groups are included in Appendix D. Heritage Preservation contracted with Aeffect, Inc. to conduct extensive pretesting for the development and administration of the HHI 2004 questionnaire. A summary of the pretests are presented in Appendix G.

The working groups recommended that the HHI 2004 include small institutions in the survey universe since typical institutional surveys tended to capture the largest and most well-known institutions, and this was an opportunity to assess preservation issues facing small institutions. Prior to any data collection, Heritage Preservation consulted with Dr. Lee-Ann Hayek, Chief Mathematical Statistician at the National Museum for Natural History, Smithsonian Institution, on planning statistical sampling for the study. Based on an estimated universe of 34,000 collecting institutions in the United States, Heritage Preservation was advised to include a total of 15,000 institutions in the study sample in order to obtain at least 5,000 responding institutions (33% response rate). The goal was to be able to report statistically accurate findings at the national level, as well as aggregated by three institutional sizes, five types of institutions, and six geographic regions. Obtaining at least 5,000 institutional replies would generally allow for data findings with 95 percent confidence level for all institutions no greater than +/- 1.3 and regional findings around +/- 3.1 percentage points.

In 2004 RMC Research Corporation was contracted by Heritage Preservation to manage the HHI 2004 study, including executing an institutional sampling plan, conducting the data collection, and analysis of data. Prior to sampling, Heritage Preservation established the study universe, a long and tedious process because one single source list did not exist. Heritage Preservation obtained comprehensive mailing lists available through directories and professional associations in 2004. Lists were culled to remove duplicates based on name of institution and zip code. Institutions not eligible for HHI 2004, such as for-profit organizations and international institutions were also excluded. A listing of the sources used to identify the universe is in Appendix H. From the estimated universe of 34,000 institutions, Heritage Preservation identified approximately 500 large institutions with a highly significant number of collections. This included all state libraries, museums, archives, and historical societies as well as major federal collections institutions. This group was referred to as Group 1. A second list of approximately 900 institutions was identified for the study which included institutions, primarily academic libraries and museums, with significant, but mid-sized, collections. This group was referred to as Group 2. The remaining 32,600 institutions with small collections were grouped by institution type and a sampling plan was developed. This sample and the institutions from Group 1 and 2 would total the desired 15,000 for inclusion in the study. Therefore, the target number for this group was 13,600 study institutions. All small archives and archaeological repositories/scientific research programs were included in the sample due to the relatively small number. Libraries, historic societies, and museums were randomly sampled within institution type and stratified by location (based on zip code). The number of sampled institutions was based on the proportion of institutions to the overall group universe. This group of study

institutions was referred to as Group 3. See Appendix I for the summary table of universe, sampling rates and sampled study institutions by sample group and institution type for HHI 2004.

Heritage Health Information 2014 Survey: Universe.
Working on behalf of IMLS, Heritage Preservation has conducted an extensive review of the source lists that were used in HHI 2004 with the purpose to identify which lists were currently active and which were updated. Many of the directories used in 2004 were duplicative of primary sources, such as the American Library Directory and the Official Museum Directory. Directories that were current and extensive were compiled as the basis for the 2014 universe. Where needed, Heritage Preservation staff called and verified the collecting procedures for directories with staff at the federal agency, private company, or association responsible for the list. Once the 2014 universe was compiled, it was compared to the 2004 universe to check for institutions that were previously included but may not have appeared in the 2014 list. This check revealed that there were no institutions from the 2004 universe that were omitted from the 2014 list. It is possible that there may be some institutions that are not in the 2014 universe, but after the extensive review of current directories and comparison to the prior universe, we believe the likelihood of institutions missing from the universe is low. IMLS and Heritage Preservation feel confident that there are no other lists that could be consulted to identify whether institutions are missing from the 2014 universe, and it is unlikely that more institutions exist. Additional searching would be an unnecessary burden of time and money on the part of IMLS and Heritage Preservation, and it would not likely yield additional results.

Records for the 2014 universe were de-duplicated based on institution name and zip code. For a description of collecting procedures and current directories, see Appendix J.
After compiling the current directories, the total universe of collecting institutions is 38,726. The sample for this survey will be 11,856 institutions. The sample size is based on obtaining a sufficient number of responding institutions in order to report statistically accurate findings at the national level, as well as aggregated by three institutional sizes, five major types of institutions, and six geographic regions.

In 2004 Heritage Preservation adapted definitions of institutional size (small, medium, large) from professional associations' publications, surveys, and were reviewed and approved by IMLS staff and the HHI 2004 study project advisors. These definitions take into account the size of the collection(s), annual budgets, size of population served, or number of staff depending on the specific type of institution. Definitions for institutional size are in Appendix C. In 2014, after updating the source lists for the universe, institutions were grouped according to the standards for large, medium, and small in each type. The institutions in Group 1 are defined as large; institutions in Group 2 are defined as medium; and institutions in Group 3 are mostly small institutions. In addition to size, significance of the collections items was also considered. There were some cases in which the institution has a medium-sized collection, but does not have particularly rare items, so these institutions have been placed in Group 3. Only 441 cases met this criterion. All other institutions in Group 3 are considered to be small size collection institutions.

Heritage Health Information 2014 Survey: Sample Design.

The sampling plan for HHI 2014 will include selective sampling and stratified random sampling by (a) surveying all large and medium institutions with significant collection holdings (Group 1 and Group 2), (b) surveying all archives and archaeological repositories/science research organizations, regardless of collection size, due to the small number of institutions (an over-sampling strategy); and (c) selecting a stratified random sample of libraries, historical societies, and museums with small collections within each type of institution and stratified by location (based on zip code). The sampling plan will result in a sample of 11,856 institutions. The distribution of institutions in the sample by Group and Type is in **Table 1**.

The sampling plan has been developed to have sufficient power for estimating survey results within a margin of error of 3.5 percentage points within five institution types (libraries, museums, historical societies, archives, and scientific research organizations) and six regional locations (New England, Mid-Atlantic, South East, Mid-West, Mountain Plains, and West).

Table 1. Universe Numbers, Sampling Rates, and Sample Sizes for HHI 2014

	Archives	Libraries	Historical Societies/ Sites	Museums	Archaeological Repositories/ Scientific Research	TOTAL
Group 1						
Sampling Rate	100%	100%	100%	100%	100%	
n=	70	179	39	163	35	486
Group 2						
Sampling Rate	100%	100%	100%	100%	100%	
n=	42	449	16	333	60	900
Group 3						
Estimated Universe n=	430	14,010	4,000	17,000	1,900	37,340
Sampling Rate	100%	27%	42%	19%	73%	
n=	430	3,734	1,681	3,238	1,387	10,470
TOTAL SAMPLE						
n=	542	4,362	1,736	3,734	1,482	11,856

On behalf of IMLS, Heritage Preservation has contracted RMC Research Corporation (RMC) to execute the sampling plan, manage the data collection, and conduct the analysis for the 2014 study. Furthermore, RMC has subcontracted with RKM Research and Communication (RKM), a research firm with computer-assisted telephone interviewing capacity, to conduct the pre-survey verification process.

Heritage Preservation will provide the universe list to RMC, on which each record will be identified by sample group and type of institution. After the Group 3 sample is selected, a verification process will be conducted using the Pre-Survey Verification phone script (Appendix E). RMC will provide the list with Group 3 institutions to RKM Research and Communication, which will verify the Group 3 institutions. Details about the verification activity are presented in Section B.2.2. This effort will also include verifying that the institution does hold collections, and it is in operation. Based on the previous HHI 2004 study, we anticipate approximately 3.5% of the sampled Group 3 institutions to be non-eligible for the study and will be removed from the sample group. The eligibility findings for each of the five institutional types for the Group 3 will be used to revise "estimated eligible sample" and universe counts. For example, if 5% of the historical societies/sites for Group 3 sample are found to be ineligible (e.g., institution has no collections or is out of operation) the estimated sample size and universe of historical societies/sites will be reduced accordingly. Groups 1 and 2 include institutions for which contact information and eligibility are already confirmed through directory listings.

This survey data will be primarily collected using a web-based online platform, with the option of a paper questionnaire for institutions without Internet access or those preferring survey submission on paper. Detailed information is presented in Section B2.

B.1.2. Response Rates

The HHI 2014 study expects an overall response rate of at least 29% of institutions submitting completed surveys. Based on the previous study, response rates are expected to vary depending on the sample group (Appendix K). Attention and awareness of preservation needs within collecting institutions has increased over the past decade, which suggests that there may be increased interest in the field. We anticipate a response rate of 70% for large institutions with significant collections (Group 1), a rate of 50% for mid-sized collection institutions (Group 2), and a rate of 40% for small size collection institutions (Group 3).

The response rate will be based on the number of completed surveys and the revised estimate of the "eligible" sample size:

> <u># of institutions with completed surveys</u>
> # of estimated eligible institutions in the sample

B.1.3. Minimum Sample Size and Accuracy

To describe the precision of institutional estimates from the HHI 2014 survey, a 95% confidence interval for estimates of universe proportions will be calculated by adding and subtracting the following formula from the sample respondent estimate:

$$1.96 \times \left(\sqrt{\left(\frac{.25}{n} \right)} \times \left(\frac{N-n}{N-1} \right) \right)$$

After data collection is completed, actual overall margin of error will be calculated. In addition to the 95% confidence interval for estimates based on all institutions, subgroup estimates will be calculated.

Taking into consideration the study sample sizes of the selective sampling groups (Group 1 and 2) and stratified random sampled group (Group 3) in addition to the varied expected response rates, an overall margin of error is estimated to be +/- 1.4. Generally, the margin of error for the sample subgroups ranges from +/- 1.0 to 3.3. The accuracy in which projections will be made for each institutional type is estimated as follows: libraries +/- 1.7; museums +/- 1.8; historical societies/sites +/- 2.8; archaeological repositories/scientific research organizations +/- 3.0; and archives =+/- 4.7. The estimated universe size, sample size, number of respondents, and estimated margin of error at the 95% confidence level for sample groups, institutional types, subgroups within sample group and type, and overall are presented in Table 2 below.

Table 2. Universe, Sample Size, Number of Respondents, and Margin of Error by Sample Group, Institutional Type, and Overall

	Archives	Libraries	Historical Societies/ Sites	Museums	Archaeological Repositories/ Scientific Research	TOTAL
Group 1						
Universe + Sample Size	70	179	39	163	35	486
Response Rate 70%	49	125	27	114	25	340
Margin of Error	7.7	4.8	10.6	5.1	10.6	2.9
Group 2						
Universe + Sample Size	42	449	16	333	60	900
Response Rate 50%	21	225	8	167	30	450
Margin of Error	15.3	4.6	25.3	5.4	12.8	3.3
Group 3						
Estimated	430	14,010	4,000	17,000	1,900	37,340

Universe						
Sample Size	430	3,734	1,681	3,238	1,387	10,470
Response Rate 40%	172	1,494	672	1,295	555	4,188
Margin of Error	5.8	2.0	2.9	2.1	3.2	1.2
TOTAL SAMPLE						
Estimated Universe	542	14,638	4,055	17,496	1,995	38,726
Sample	542	4,084	1,716	3,519	1,442	11,303
Number of Respondents	242	1,844	707	1,576	610	4,979
Response Rate	45%	45%	41%	45%	42%	44%
Margin of Error	4.7	1.7	2.8	1.8	3.0	1.0

B.1.4. Sample Unit and Non-response Item Weight Development

An overall goal of the HHI 2014 study is to project findings to the estimated universe of institutions holding collections in the United States as well as the condition of the number of collection items. The development of sample weights at the institution level and data imputation for non-response questions will allow for estimated projections to be made.

To adjust for non-response and random sampling respondent weights will be developed:
- Non-response for Group 1, 2 and 3
     Weight=reciprocal of the response rate for the institution type

- Random sample stratification for Group 3
     Weight=reciprocal of its probability of selection for the institution type

For Group 3, the product of the above adjustments will be the final analytic weight. Weights will be calculated for each survey respondent and applied to all of the respondent's survey responses.

Item non-response will also be addressed within the study. Questions regarding the number of specific collection items or the percent of collection items in urgent need of care may not be known to respondents. Respondents will be encouraged to estimate figures, but also will be provided the option to respond "don't know" to these items. One of the main objectives of the HHI 2014 is to report on the conditions of all collection items in the United States. Therefore, missing data will be imputed with values from similar institutions. For each of the individual types of collections, only those institutions holding that collection will be identified and aggregated by institution size (large, medium, small) and specific type of institution (i.e.

archives, public libraries, special libraries, academic libraries, independent research libraries, historical societies, art museums, history museums/sites, science museums, and archaeological repositories/scientific research organizations). The *median* quantity of collection items and *mean* preservation conditions will be calculated for each possible subgroup. If a subgroup contains fewer than 10 institutions contributing to the mean or median, that subgroup will be combined with another subgroup of the same size and similar type.

Software. The software package IBM Statistical Package for the Social Sciences (SPSS) will be used for computing and implementing sample weights and data to be imputed for unit non-response. Analyses, such as cross-tabulations (i.e., by type, institution size, region) and quantitative totals will be conducted for unweighted and weighted data.

B.2.   Procedures for the Collection of Information

After OMB approval is received and the Group 3 sample verification contact is completed, RMC will begin field operations. All potential study participants will be contacted via email or postal mail explaining the importance of the study, asking for participation, and providing instructions on how to log on to the study website to access the survey or how to submit their responses by postal mail. In addition, respondents will have access to online help if terms need to be defined. The paper questionnaire will include a glossary of terms and contact information for respondents who need help. Several follow-up emails will be sent as a reminder to non-respondents and respondents who have only partially completed the survey. The online survey development, validating the sample and contact information, emailing, and tracking are described below.

B.2.1. Online Survey Development

The online survey collection mechanism that will be employed for HHI 2014 will be built upon an existing application developed by Community Logic, Inc., a subcontractor to RMC Research. This application has been developed with the Microsoft web technology stack, employing ASP.NET 4 with SQL Server 2005 as the backend data repository. The mechanism employs the jQuery JavaScript library and the jQuery UI plugin library to support cross-browser compatibility and section 508 compliance. Community Logic, Inc. will work with RMC's IT Director, to install, test, and administer the survey mechanism on RMC's web server.

The HHI 2014 questionnaire will be loaded into Community Logic's existing survey engine mechanism. This mechanism stores forms as objects that define the field types, labels, and field-level validation checks used for a particular form. Because the structure of the survey instrument is stored as a data object, it can be easily modified to add the additional sections and questions that are envisioned for HHI 2014. This will allow us to efficiently modify the existing instrument in collaboration with Heritage Preservation within the proposed time schedule in response to any issues that arise in pre-testing.

Data validation will be performed on both the respondent-side and server-side. Respondent-side validation will be handled with the JQuery Validation plugin developed by Jörn Zaefferer, supplemented with a set of custom JavaScript validation routines to validate relationships

between field responses. All validation routines will be mirrored on the server side to validate data before it is accepted for submission. See Appendix F for screenshots of the web survey.

The existing survey engine mechanism provides all of the field-level validation requirements of the study. These include data type validation for numeric, date, four-digit year, currency, percent, email, and URL fields. Numeric, date, and single- and multiple-choice fields provide both range validation and comparison testing against other discrete fields in any section of the form for internal consistency checks. Automatically calculated fields can be added to perform mathematic operations against any number of discrete fields in any section of the form. Individual fields can be declared as required, and the required flag can respond to customized skip logic against answers provided in other fields. The survey mechanism provides immediate feedback response to any validation errors on the respondent-side when a response fails an edit, and performs server-side validation checks to ensure data integrity. Customized pop-up help can be provided for any individual field or section.

B.2.2. Sample and Contact Validation, Emailing and Tracking

RKM Research and Communication will verify the Group 3 listings. RMC will provide the sample list of Group 3 institutions to RKM Research and Communication. RKM interviewers will use this list to call these institutions in order to introduce the study to these institutions and to encourage participation. Interviewers will also verify contact information, including names and email addresses. All RKM interviewers will undergo training and orientation, regardless of their level of experience, prior to making phone calls. Institutional contacts seeking more information about the HHI 2014 study will be referred to the staff contact at IMLS and the Project Director at Heritage Preservation. The disposition of all phone call results will be recorded for each institution, such as numbers out of operation, duplicate records, institutions that do not meet inclusion criteria, and institutions that request not to participate in the study. The telephone interviewer script is in Appendix E.

The verification call serves three key purposes: 1) to validate the eligibility of sampled institutions; 2) to assess the Internet capability of institutions; and 3) to identify the most appropriate contact within the institution to direct the upcoming correspondence about the study. In particular, it is important to validate that each of the sampled Group 3 institutions actually hold collections and is a non-profit institution, as this influences the universe of eligible institutions and affects the sampling frame. Just as important is the identification of those institutions not eligible. The non-eligible records will be analyzed by institution type and region in order to make adjustments to the estimated population, which is based on several directories and professional association listings. Ultimately survey results will be projected to the national level by type and region. Therefore it is critical to verify the state of each institution, whether eligible or not, during the validation call. We propose asking all institutions in the sample group after validating the name and before the eligibility screening. Through the branching, the verification of the mailing address will only be asked of those institutions that indicate that either they do not have Internet access or that they prefer a paper submission process.

RMC will load the validated institutional contact information (institution name, address, contact name, email address) into the data collection tool with the survey engine's existing import mechanism, which will generate a unique institution-level username and password to serve as a unique identifying pattern. In addition to the generated username and password, study invitations, instructions, and a link to the online survey will be distributed to the study sample via email or postal mail.

A unique email address will be established to send and receive emails to respondents. Bounced and non-deliverable email addresses will be monitored and tracked. RMC will provide IMLS and Heritage Preservation with any invalid email addresses for follow-up. When corrected email addresses are obtained, RMC will resend study instructions and link to the online survey. We expect the overall data collection period to last eight weeks from late September to middle November, 2014.

B.3.   Methods to Secure Cooperation, Maximize Response Rates, and Deal with Non-Response

B.3.1. Gaining Cooperation

In addition to phone verification of the sample (described above in B.2.2) notifying participants of the survey, extensive publicity is planned for announcing the HHI 2014 study. Heritage Preservation, including their staff and Board of Directors, has asked professional associations to inform their constituents about the survey and to publicize the survey to encourage study participation. The Heritage Preservation staff and Board are comprised of well-known leaders in the collections care and conservation communities. Supportive professional associations in the cultural heritage industry will be listed on the survey website to demonstrate their support of the project.

Heritage Preservation will publicize the HHI 2014 through press releases distributed through the IMLS press list of professional archive, library, and museum associations and publications. Plans also include announcements in various professional newsletters, listservs, and at association meetings, including the American Association of State and Local History (AASLH), the American Alliance of Museums (AAM), the American Institute of Conservation (AIC), Small Museums Association (SMA), and the American Libraries Association (ALA). Publicity at professional association conferences and through IMLS's press contacts will reach all institutions in Groups 1 & 2 and some in Group 3. Additional efforts to reach Group 3 will include contacting potential respondents directly through listservs and Connecting 2 Collections (C2C) community announcements. Some Group 3 participants may not see the study publicized at national association conferences or on professional listservs, which emphasizes the importance of the pre-survey verification phone call to notify them of the study. Where possible, efforts will also be made to attend regional and state conferences, such as the Mid-Atlantic Association of Museums and the Virginia Museums Association to reach Group 3 institutions. Study announcements and pamphlets distributed at conferences will explain how the information will be used, and the importance of participating in the study.

Heritage Preservation has been, and will continue to, advertise the study on its primary and affiliated websites. Once data collection begins, Heritage Preservation will email or mail an invitation to organizations in the sample notifying them of their inclusion in the sample, the importance of the study for the cultural heritage field, and how the results will help them manage and monitor their collections care practices. An email announcement will also be sent to each participant announcing the release of the final results from the survey.

B.3.2. Methods to Maximize Response Rates

For the HHI 2014 study, it is anticipated that at least an overall response rate of 29% can be reached. We will employ a number of techniques to enhance the response rates for completed surveys, described below.

All Participants will receive the following:

Survey Support. In addition to the field-level online help provided by the survey website, technical support will be provided by email or phone by RMC and content support will be provided by email or phone by Heritage Preservation.

Online survey and multiple sessions allowing multiple respondents. Study participants will be encouraged to log into the website and download a PDF version of the survey, instructions, and reference guide before entering any institutional collections information. By reviewing the survey, participants are able to manage the data collection effort, such as how long it will take to gather the information and if more than one staff member is needed to complete the survey data.

Once the hard copy survey is completed, participants are encouraged to enter the information into the online web-based survey. As explained below, multiple online respondents are allowed over multiple sessions if this is more appropriate for the institution. For those participants who prefer not to enter their information online, they will be asked to mail the completed hard copy survey to RMC Research, where staff will then enter the data from printed mailed submissions via the website mechanism.

We will employ a two-tier security model to ensure confidentiality and data security. Once a user has entered the website with their institution's username and password, they will be asked to establish an individual user account with an individual username and password that is automatically tied to their institution. This process will allow multiple staff members from the same institution to complete sections of the survey instrument over separate sessions, while providing individual user authentication and tracking of responses.

Utility. As stated by previous HHI 2004 respondents, the effort put into completing the survey and having a hard copy record of the institution's collections status is highly beneficial for sharing results with stakeholders and funders. Respondents will be notified via email when the HHI 2014 study results are posted on the HP website so institutions will be able to make comparisons between their preservation activities and conditions and those of other similar institutional type, size and regional location.

Skip routines. Data quality will be enhanced not only through field-level data validation, but through a set of user feedback mechanisms. These will include logical relationship validation checks, pre-fills, and skip logic. For example, a user who responds "No" to the question "Does your institution have a written emergency/disaster plan that includes the collection" should logically respond "No" to the follow up question "Is your staff trained to carry it out?" Using custom JavaScript functions that employ the jQuery library for cross-browser functionality, we will conduct cross-field logical relationship checks such as this, and pre-select the logical response based on user input. In addition to logical pre-fills, user feedback in the form of pop-up warnings will be triggered if the user makes a logically improbable choice. In consultation with Heritage Preservation, we will determine what responses should trigger skip logic to disable questions that the user need not answer. In addition, we will provide a set of respondent-side tools to guide the user in entering valid information. For example, we will provide a pop-up FTE calculator to help ensure that the interpretation of FTEs is consistent across survey responses.

Completion bar. The survey will be presented in section-level and/or sub-section level pages, as desired, with data validation and response storage occurring on each individual page. Survey respondents will be able to save their responses at any time and return to complete the survey over multiple visits. For the online survey, a visual feedback mechanism in the form of a "subway map" that indicates at a glance which survey sections had been completed and which had not will be developed. Each node of the map will link to that particular section of the survey. In addition, we will add textual support for ADA compliance. A final submission check will be conducted before survey completion, and any missing information will be highlighted for the user.

Interactive charts and maps. Respondents will be offered an additional incentive to complete their surveys. Upon completion, respondents will have access to a dashboard page that will provide current aggregate information about survey responses.   Aggregate information such as current response rate by institution type and geography will be offered in the form of a set of interactive charts and maps. We hope to find this mechanism helpful in stimulating participation among peer institutions.

Confidentiality and data security. As one measure to protect the confidentiality of the data under the requirements of the Privacy Act of 1974 and the E-Government Act of 2002, data transferred to and from the survey website will take place over an SSL connection.

RMC Research Corporation's IT Director performs an annual security self-assessment review patterned on the NIST guidelines for low-baseline security assessments following NIST Special Publication 800-53. An Information Assurance Independent Verification and Validation Management Committee (IV&V MC) conducted a review of existing findings relevant on behalf of the low impact system OCO/NCLB Blue Ribbon System Certification in which RMC Research achieved accreditation through June 2012. RMC maintains a comprehensive System Security Plan (SSP) that is updated annually in accordance with NIST guidelines. System Security Plan requirements include: the updating of all applicable security controls; continuous monitoring; 3rd

party vendor security considerations; authorizing connections; physical security considerations; the creation of off-site backups; and the purging of retired equipment.

Response rate monitoring and follow-up reminders. RMC will monitor response rates on a weekly basis and submit status reports to Heritage Preservation. Response rates will be analyzed for each sample group, institution type, and region to ensure adequate sample representation. In addition, rates will be segregated by survey completion and partial completion. Over the eight week data collection period, reminder emails including survey information and webpage links will be sent to non-responders and partial responders during week 4 and week 7. In addition, RMC plans to send an oversized postcard to non-respondents during week 4 via postal mail directed to the contact person to encourage participation. In the event response rates are lower than expected, Heritage Preservation may decide to extend the data collection period by two weeks and RMC will send out another email during week 8 announcing the extension and asking for participation.

Group 3 Recruitment/Follow-up Strategies:

The following steps will be taken to encourage study participation for Sample Group 3. This includes a pre-notification to alert institutions to the survey and follow-up measures to non-responders to bolster response rates.
- Pre-notification and verification process via phone call (n=10,470). The protocol script is attached (Appendix E).
    - o Introduce HHI study, importance of participation, and notification of sample selection
    - o Verify/obtain contact information including institution name, address, appropriate contact name, position
    - o Verify/obtain access to internet, appropriate email address
    - o Obtain preference for survey mode (email, mail)
- Survey Deployed with Field Period of 8 weeks
    - o Mail survey package to institutions preferring paper survey and those without email addresses on file
    - o Email survey package to institutions preferring Internet participation
- Follow-up Reminders
    - o At four weeks in the field,
        - ▪ Email reminder sent to all non-respondents with email addresses
    - o At five weeks in the field,
        - ▪ Assess response rates for each type of institution
        - ▪ Follow-up reminder phone calls to archives (85% follow-up, n=365). See HHI 2014 Follow-Up Phone Call script in Appendix M.
    - o At six weeks in the field,
        - ▪ Oversized postcard sent via postal mail to all non-respondents. See HHI 2014 Follow-Up Postcards in Appendix N.
    - o At six weeks in the field,

- Assuming a 20% response rate to date, all non-respondents for historical societies and scientific research organizations contacted by phone (n=2,455).
- Reminder phone call. See HHI 2014 Follow-Up Phone Call script in Appendix M.
- Email survey package again
  - Resend survey packages as requested from follow-up phone calls
  - During week 8 of field period
    - Announce survey extension (email and mail)
    - Second follow-up calls to non-responding archives (60% follow-up, n=258) See HHI 2014 Follow-Up Phone Call script in Appendix M.
  - During week 10 (field period extension)
    - Final reminder email to all non-respondents

The following timeline (**Figure 1**) presents when and methods of how institutions will be contacted throughout the study. Specific institution types will receive additional follow-ups to increase response rates, as described.

Figure 1. Timeline of strategies to increase response rate in small institutions (Group 3)

| | | | Survey Field Period (weeks 1 - 8) | | | | | | | | Survey Extention (weeks 9-11) | | | | |
	Verification/Notification	Survey Deployed	week 1	week 2	week 3	week 4	week 5	week 6	week 7	week 8	week 9	week 10	week 11		
Phone Call	ALL														
Survey Package (mail/email)		ALL													
Email Reminders					ALL			ALL	ALL		ALL				
Postcard Reminders							ALL								
Follow-up Phone						1	2,3		1						

1 = Archives only
2 = Historical Societies
3 = Scientific Research Organizations

B.3.3 Statistical Approaches to Nonresponse Bias

Based on the 90% response rate obtained in the 2004 study, we can assume a response rate higher than 80% for the Group 1 institutions, so a non-response bias study for this group is not being anticipated. However, in the event the response rate for this group is lower than 80% we will conduct a non-response bias study similar to the study for Group 2 institutions described below.

If the total number of Group 2 institution is approximately 900, based on the same response rate as in 2004, we can expect 405 respondents (45%). Based on the 2004 finding that all Group 2 institutions had internet access 10 years ago, all 495 non-respondents will be contacted via email requesting their participation in a very brief online survey (HHI 2014 Non-Response Bias Questionnaire). The survey questions will be embedded in the email with the online survey link so recipients will know ahead of time what the questions are and see the time it will take to participate. Comparisons will be made between the respondents to the original survey and non-respondents that complete the follow-up survey. Obtaining a response rate of 20% (n=99) will result in enough power to detect any differences.

We assume an overall response rate of 40% for the Group 3 institutions, which will result in 6,282 non-respondents. We will draw a stratified sample of half of the non-respondents (n=3,141) based on type of institution. Our goal for the follow-up survey will be to obtain enough responses to support a margin of error between 4.3 and 4.9. Table 3 shows that this will require a response rate of 12-15% of the 3,141 sampled non-respondents. We will use non-overlapping confidence intervals as an indicator of a significant difference between respondents and non-respondents that respond to the follow-up survey. With a margin of error of 3.5 for the original survey and 4.3 to 4.9 for the follow-up survey, any difference larger than 7.8 to 8.4 will indicate a significant difference between the original respondents and non-respondents, indicating the potential for bias due to non-response in the original survey results.

Table 3. Group 3 Sample needed for Non-Response Bias Follow-up Study

Survey Sample size	10,470
Est. Number of Responses at 40%	4,188
Number of Non-respondents	6,282
Send Non-Respondent Survey to Sample	3,141
Response Rate of 12% - 15%	377-471

From the notification/verification phase of the HHI 2014 study, we will know the status of internet access for each sampled institution for Group 3. For institutions that did not respond to the original survey which have indicated internet access, we will send them an email request to complete a brief online follow-up survey (Non-Response Bias Questionnaire). The survey questions will be embedded in the email with the online survey link so recipients will know ahead of time what the questions are and see how little time it will take to participate. For non-responding institutions not having internet access and/or requested to have a paper survey, a paper version of the non-response questionnaire will be mailed to the institutions (HHI 2014 NonRespBias Questionnaire Paper version). Comparisons will be made between the respondents to the original survey and non-respondents that complete the follow-up survey.

During the notification and verification phase of the study conducted by phone, all institutions will be asked about Internet access. Response mode will be included on the data file. Analyses will be conducted to examine response as a function of Internet capability. It will also be examined in the non-response bias analysis to determine whether it was a factor in possible response bias.

B.4.   Tests to Minimize Burden

Background. As stated in Section B1.1, IMLS partnered with Heritage Preservation to develop and conduct the first HHI project starting in 2001. Heritage Preservation established an Institutional Advisory Committee of 35 professional associations and federal agencies representing collecting institutions and 66 working group members (see Appendix D) to advise on the survey development and implementation of the HHI 2014. Because advisory and working group members represented archives, libraries, historical societies, museums, and scientific research organizations, Heritage Preservation was able to build a consensus on neutral terminology and avoided technical language and jargon to ensure study participants of any professional level would understand the questions.

In the summer and fall of 2002, a survey research firm Aeffect Inc., of Deerfield, Illinois was contracted to pretest the survey instrument and administration. The first pretest focused on the survey questions and 30 volunteers participated and of those, 18 in-depth interviews were conducted to obtain feedback on the clarity of the survey questions and response options. Accordingly, adjustments were made to the survey and a second pretest was conducted. A random sample of 202 institutions across the country were asked to participate in the second pretest resulting in 75 additional institutions providing feedback and suggestions further refining the survey instrument, term definitions, and survey administration. Appendix G presents a summary of both pretests.

Heritage Health Information 2014.
The majority of questions and response options for the current study will remain the same as the first study except some response options are combined due to findings from the first study. However, collection-holding institutions have experienced a huge increase in digitizing their collections over the past ten years. To accommodate this change, Heritage Preservation has worked with a newly formed digital advisory group to formulate appropriate questions about these digital functions. Group members are from the Office of Preservation and Access at National Endowment for the Humanities, Office of Library Services and Office of Museum Services at IMLS. The HHI 2014 survey will incorporate these new survey questions after a pretest process. Questions include the possible functions an institution carries out (i.e., digitization, creation of metadata, format migration, tool development, and redundancy), staffing responsibilities, existence of policies, professional development, and conditions of digital materials, such as images, texts, software, and games. Similar to the HHI 2004 survey, a glossary of terms will be available. The pretest validated its usefulness and clarity. Representatives from 9 institutions which had digital holdings and were highly experienced in digitizing collections were asked to provide feedback on the survey items in terms of clarity and comprehensiveness, including the new survey questions, response options, and glossary; no data were collected. All feedback from pretest participants was shared with IMLS, Heritage

Preservation, and the digital advisory group members to finalize the current survey questions on digital collections. The findings informed the revision of the questions, which are included in the fully revised 2014 web and paper questionnaires. See the Appendix O for the results of the pretested digital questions.

Aside from pretesting the survey questions, prior to any data collection, pretesting the HHI 2014 online survey mechanism is planned. The testing will assess the survey validation routines; skip patterns, and overall time it takes to enter survey data, as well as specific timing for sections of the questionnaire. There will be two phases of technical pretesting. The first phase will include internal tests done by staff from RMC, Heritage Preservation, and IMLS. Once the first phase is completed and needed mechanism adjustments are made, colleagues from cultural agencies, such as the National Park Service, National Endowment for the Humanities, Library of Congress, and National Archive Administration will be asked to "test drive the survey mechanism with dummy data". All adjustments to the web-based survey will be made prior to any study data collection.

B.5. Individuals Responsible for Study Design and Performance

The following individuals are responsible for the study design and the collection and analysis of the data for Heritage Health Information 2014 Survey.

Personnel Involved with HHI 2014

Person	Address	Email / Phone
Institute for Museum and Library Services **Christopher J. Reich** Senior Advisor, Office of Museum Services	1800 M Street NW 9th Floor Washington, DC 20036-5802	creich@imls.gov 202-653-4685
Deanne Swan, Ph.D. Senior Statistician	1800 M Street NW 9th Floor Washington, DC 20036-5802	dswan@imls.gov 202-653-4759
Heritage Preservation **Lesley Langa** Director, HHI 2014	1012 14th Street NW Suite 1200 Washington, DC 20005	llanga@heritagepreservation.org 202-233-0824
RMC Research Corp. (RMC) **Kim Streitburger** Project Manager	1000 Market St. Bldg. 2 Portsmouth, NH 03801	kstreitburger@rmcres.com 800-258-0802
Allen Schenck, Ph.D. Senior Research Associate	1501 Wilson Blvd. Arlington, VA 22209	schencka@rmcarl.com 703-558-4912
Tracey Martin, Ph.D. Research Associate	1000 Market St. Bldg. 2 Portsmouth, NH 03801	tmartin@rmcres.com 800-258-0802
Tim Golden IT Director	1000 Market St. Bldg. 2 Portsmouth, NH 03801	tgolden@rmcres.com 800-258-0802
Community Logic Inc. **Doug DeNatale** Web Developer	166 Hawthorne Street Malden, MA 02148	denatale@mediaone.net 781-956-5682
RKM Research and Communications **R. Kelly Myers** President Project Manager-Phone Verification	1039 Islington St. Portsmouth, NH 03801	kmyers@rkm-research.com 603-319-4269