

## **OMB Supporting Statement B**

### **IMLS Study of Sustainability of Digitized Special Collections**

#### **Part B: Collections of Information Employing Statistical Methods**

##### **B1 Potential Respondent Sampling and Selection Methods**

The study consists of case studies of eight individual digitized special collections projects.

##### ***Selection of the case study subjects***

This study will identify and examine 8 case studies of a specific type of digital resource, digitized special collections.<sup>1</sup> The Ithaca S+R and ARL research team will identify examples of digitized special collections that show evidence of being sustainability “success stories”, that is, projects that demonstrate the elements we define as constituting a sustainable project. This approach supports the ultimate aim of this work: to provide useful models for the community to observe, emulate, challenge, and improve upon. By starting with “success” cases, we hope to be able to better understand the process taken to achieve these outcomes. The selection methodology outlined below may also allow us to start to see how the paths to sustainability might be different, given different institutional settings, or the type of initial funding they receive.

Perhaps most important, we hope that by selecting cases in a range of different institutional contexts, readers will easily find examples of cases that are most relevant to them, in terms of the context in which they exist; The more we can offer cases that our readers feel “look like me”, the easier it may be for them to consider applying some of the strategies described in the cases they read to their own collections.

Selection Criteria An initial screening will seek to identify certain characteristics which taken together describe a project that is sustainable, and will also ensure that the projects studied represent a wide range of institutional contexts. The screening protocol will be finalized by the project team and will be reviewed by the expert advisory board of the project, which includes prominent university librarians, museum directors, and other representative professionals. Once reviewed by the advisory team, the protocol will be submitted to IMLS for final review and approval before case screening begins. A draft screening protocol is attached. While some selection criteria will be possible to identify through desk research, given the lack of independent sources for data for monitoring project sustainability, we will rely heavily on answers provided by potential respondents. Wherever possible, data will be checked against independent sources. For example, the date a web resource launches is sometimes available as part of the history provided on the website; where a project was grant-funded, dates of the grant can be verified independently through primary observation. In other cases, however, we will need to rely on the candor, accuracy and honesty of the project leaders. For example, there is no way to independently verify self-reports of user base. However, to assess the validity of claims, the team will ask a set of questions to determine how the user statistics were derived. Respondents will be asked how they came to the user figure, which web analytics tools were used, and how long the metrics have been collected, etc.

---

<sup>1</sup> These are collections of rare or unique content that libraries or museums have chosen to digitize.

Screening protocol. In a first phase of desk research, digitized special collections will be segmented according to type of institution, budget size, and source of initial funding. (For further detail, see the Draft Desk Research Protocol on “segmentation”) After we have classified projects according to institution type, size, and source of funding, a further screening via desk research will allow us to weed out ineligible projects and to prioritize others.

Projects will be screened according to the three primary screening criteria identified for this study, listed below. Desk research will supply the information about the projects that we will need in order to determine whether or not they fit these criteria, and the results will be entered into a spreadsheet:

Longevity - projects in existence for more than 2 years past launch

- Here, “launch” will mean that time at which the content becomes accessible to target users.
- We will look on the websites of projects for evidence of their start dates, either on their homepages or in a section on their histories. If this is unsuccessful, attempts will be made to determine the start dates by searching the web and looking for press releases or other articles with these details. Results will be verified by looking at grant reports, if they are available. Whether or not projects have been in existence for more than two years will be indicated by either “yes” or “no.”

Financial Stability – projects that are able to cover operating costs (or better) in a reliable manner over time

- This will be determined by asking if revenues meet or exceed expenses for the project. We know from past research that project leaders may not have at hand an easy answer to this question. Other ways we can address this issue include asking how reliable the funding is to support the resource, if the resource has sufficient support to stay current or consider upgrades (or conversely, how it has dealt with reducing functionality or activities), and if the project team has attempted (and/or succeeded in implementing) any innovative revenue strategies
- We will determine financial stability by looking for evidence of creative and varied attempts to develop funding streams, including efforts to generate revenue (advertising, requests for donations, evidence of sponsors, pay models, etc.) on the projects’ sites or in other related locations (e.g., the host’s site). Some may also include this information in their project histories. If the projects are still active, whether or not the sites are current and have been updated recently and regularly will also serve as an indicator of financial stability. Sites will then be rated on a scale of 1-5, with 5 indicating evidence of a variety of robust funding streams, and 1 indicating a site that is no longer operating.

Public Benefit – projects that show evidence of a strong value to their users. This can mean showing evidence of many users, users who are deeply engaged in the resource, or other indicators, including awards, citations, or other measures.

- Numbers concerning users may be difficult to obtain, particularly if they are not being measured by the project itself. We may have to consider this element somewhat exploratory, and accept that there will be several acceptable ways to measure the strength or size of an audience. Most important, “size” of an audience is relative; 1000 devoted users of a niche product may be much more “valuable” than

1000 occasional users of a mass-market product. Furthermore, we recognize that resources may target different audiences; projects dedicated to teaching will differ from those intended for research or preservation. To operationalize this, we will develop criteria that permits project leaders to self-assess the current degree of audience usage and engagement and the degree to which it is meeting their goals/targets.

- In order to account for the different kinds of audiences projects might have, we will determine their public impact not just by looking for user comments (e.g., on project blogs) and visitor counters on the websites (which can be unreliable). We will also look for social media activity and the projects' presence on the web more generally, asking if they are being cited in academic journals, the sites of professional organizations, or in the news media. Sites that have received awards or special accolades will also be noted. They will then be rated on a scale of 1-5, with 5 signaling a project with strong public impact, whether it is for a small but active niche group or for a wide-ranging group that is less active. 1 will mean that the project appears to have had little uptake from its audience or the community in which it operates.

The desk research will help to weed out projects that are of little interest for a study of sustainability in digital archives, whether because they no longer exist or because they have been in operation for less than two years, for example. For certain criteria, including “public benefit” and “financial stability” where obtaining accurate data from desk research alone may not be possible, we will be careful to not hastily exclude cases because we do not have access to full information. Here, we will use the ratings to prioritize the strongest cases, and then a phone screen will allow us to more accurately assess the degree to which projects fit our selection criteria. The Draft Phone Screen Protocol provides further detail on the questions to be asked and the means for assessing projects for fit.

\

Next, the research team will address the institutional context of the project itself, with some cases located at libraries and some at museums, the two main audiences for this work.

Previous research has suggested that there may be different strategies at play and different types of challenges to face given certain conditions, so we will select cases that allow us to examine success cases that are:

- From Large versus Small institutions
  - This will be based on the size of operating budget of the institution, with appropriate break points set for different institution types. For example, in the case of academic libraries, the breaking point will be having an operating budget in excess of \$1million.<sup>2</sup>
- Created through external [grant] funding versus internal sources of funding
  - External grants could be from any sources, public or private
  - “Internal funding” would mean funded primarily internally – more than 50%, for example – as it will be difficult to find projects that did not benefit from any outside support at all.

---

<sup>2</sup> National Center for Education Statistics, *Academic Libraries: 2010 First Look*. Table 8 (page 11) shows that approximately 30% of all academic libraries have operating budgets in excess of \$1,000,000. in this <http://nces.ed.gov/pubs2012/2012365.pdf>

A final selection grid would look like this:

		<b>Museums/ PLs</b>	<b>Academic Libraries</b>
Large	Grant	1	1
	Internally funded	1	1
Small	Grant	1	1
	Internally funded	1	1

### Case Screening and Identification

The final case study subjects will be chosen through a process of exploration of likely targets. Previous research has helped Ithaka S+R to become familiar with those who fund and create digitized resources, and these contacts will help in developing an initial list of 16-24 potential targets. Among the sources to be used to develop this list are:

- 1.) IMLS grantees through the National Leadership Grants Program to create or enhance digitized special collections;
- 2.) Members of ARL’s Transforming Special Collections in the Digital Age Working Group, who may have recommendations;
- 3.) Program officers at the National Endowment for the Humanities, in particular through the Preservation and Access program;
- 4.) Additional desk research.

Given that all cases must meet our initial screening criteria as exemplary cases, (ie, those that “reflect strong, positive examples of the phenomenon of interest”<sup>3</sup>) requires extensive research in order to identify viable cases that meet this initial criteria.<sup>4</sup> The research team will begin with building a list of possible targets drawn from the sources above. To evaluate whether a project meets the selection criteria, the research team will examine and evaluate publicly available data through a process of desk research. Data sources will include university or library websites with known strengths in digital humanities, news sources, profiles in previous reports or case studies, among others. Individual cases will be selected to include projects that vary in their longevity, the scope of the project (based on the size of the project as well as the size of the audience), and the effectiveness and innovativeness of their sustainability models. In addition, it is likely that knowledgeable individuals – including program officers – can provide some strong leads based on their first-hand familiarity with many projects. .

Once the research team has developed this list, which will include a few choices for each desired criterion, ARL and Ithaka S+R will share recommendations for prioritization with IMLS. At that point, the research team will begin to contact the project leaders, to introduce them to the project and invite them to participate. In order to ensure that project leaders are aware of the criteria for the study, the

---

<sup>3</sup> Robert K. Yin, *Applications of Case Study Research*, second edition. Applied Social Research Methods series, vol 34, p.13. See box 4: Exemplary Case Designs. Also see Alexander L. George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences* (MIT Press, 2005).

<sup>4</sup> Yin, p. 14. Note that in one study this phase accounted for 20% of the total resources of the project.

research team will explicitly include the criteria and characteristics that will be used to determine the final selection.

## **B2 Procedures for Collecting Information**

Once the research team has identified the desired projects for the case studies, they will contact project leaders by email and by phone to invite them to participate in the study and to discuss what this will entail. Once they have agreed to participate, phone interviews will be scheduled with the primary contact to discuss the project background and to identify others working on the project who might be important to interview as well. This initial interview will provide the research team with a broad understanding of the project and frame subsequent conversations with project stakeholders; in some cases, these conversations may lead the research team to determine that the project is unsuitable for the study. Some reasons a project may be deemed unsuitable for inclusion are that the project leads choose not to participate, that the research team determines that in fact the criteria for inclusion were not met. In this event, the research team will proceed to contact the next project on the priority list.

Interviews will be conducted via telephone or in person at test sites with key people involved in creating or who are currently involved in maintaining the digitized special collections. One to four respondents will be contacted for each case study. In addition to interviews, respondents will be asked to send the research team any relevant documentation they would like to share that will help to understand key aspects of their work. Interviews will last 90 minutes each, and may involve follow up calls or emails for further clarification.

Interviews will be conducted by teams of Ithaca S+R researchers, with experience in field-based qualitative research and semi-structured interviewing of the type that will be used in this study. All researchers involved in the fieldwork will be trained with respect to the objectives of the study, study procedures and protocols, and the types of information sought for the study.

## **B3 Response Rates and Non-Responses**

Ithaca S+R will identify an initial short list of 16-24 possible case study subjects, representing 2-3 potential targets for each type we seek to represent (see selection grid on page 2). The prioritized list will be agreed with ARL and our advisory board. From this list, Ithaca S+R and ARL will select the eight projects for which we will research and write case studies.

Participation will be voluntary and based on selection criteria and on project leaders' willingness to participate and ability to schedule and accommodate site visits. Identifying likely targets in each category will require careful and thorough research. Given the extensive network of contacts, and the hundreds of projects from which the research team will choose, we do not anticipate difficulty in identifying potential targets. Since the methodology requires identifying success cases, however, the research team is prepared to screen more than 3 projects per segment, should this prove to be necessary.

## **B4 Tests of Procedures or Methods**

The project team has created an expert advisory board, including prominent university librarians, museum directors, and other representative professionals. This board has already provided feedback on the methodology and will conduct a review of both the selection criteria and the interview guides for the case studies before materials are submitted to IMLS for final review and approval and before work begins.

### **B5 Contact Information for Statistical or Design Consultants**

Staff members at ARL and Ithaka S+R are primarily responsible for conducting the case study research. The entire project team includes:

Senior Advisor: **Charles B. Lowry**, Executive Director, Association of Research Libraries

Project Lead for ARL: **Judy Ruttenberg**, Program Director for Transforming Research Libraries

Project Manager: **Nancy L. Maron**, Program Manager, Ithaka S+R

Senior Advisor: **Deanna Marcum**, Managing Director of Ithaka S+R, ITHAKA

Publication Design: **Lee Anne George**, Publications Program Officer, Association of Research Libraries

Staff: **Jason Yun and Sarah Pickle**, Analysts, Ithaka S+R

Technical Support: **Tricia Donovan**, ARL Communications and Project Coordinator, Association of Research Libraries

Federal Contact: **Carlos Manjarrez**, Director of Planning, Research, and Evaluation, Institute of Museum and Library Services