

Institute of Museum and Library Services (IMLS)
The Impact of Free Access to Computers and the Internet in Public Libraries

A. Justification

A1. Need and Legal Basis

The advent of the Internet and computer technology has radically changed the way people live. Public libraries have been at the forefront of championing digital inclusion through partnerships with the Bill & Melinda Gates Foundation, the IMLS, other national organizations, and their own communities. As a result, virtually every library in the U.S., provides public access computing to their community. This includes not only digital resources, databases, networked and virtual services, but also the training, technical assistance, and staff knowledgeable in technology services as support for the users. These technologies “range from basic services (computers, Internet, and online catalogues) to sophisticated, interconnected technologies that bring digital resources and virtual services to patrons” (IMLS, 2007)

While these services are widely available, past research has produced little evidence that shows a relationship between public access computers and community benefits. Past decision-making regarding public access computer services has been based on such measures as number of users/sessions, length of time computers are in use, number of unfilled requests, and results of satisfaction surveys (e.g., Jaeger, Bertot, & McClure, 2007). In many ways, these measures are reminiscent of those lamented by Zweizig and Dervin (1977) as being inadequate for understanding the deep impacts of library services in communities. To recognize and communicate the value of free access to public access computers in libraries as it is accrued to individuals, their families and communities, better methods that ask different questions and yield robust data are needed that will empower library decision-makers, especially in efforts to guide social policy. Along this line, research is needed that identifies the negative impact of absent or low-quality public access computer access, and most importantly, captures the correlations between specific types of library public access computers and desirable outcomes.

A2. Purpose and Use of Collected Information

The materials developed from the analyses of this data collection effort are intended to reach multiple audiences—IMLS staff, policymakers at the state and federal levels of government, library professionals, academic researchers, and others interested in understanding the impact of public access computers in libraries.

The purpose of these materials is: (a) to develop robust and broadly applicable social, economic, educational, and health indicators for the impact of free access to computers, the Internet and related services in public libraries on individuals, families and communities; and (b) to apply those indicators to validate their robustness and document positive and/or negative results from the presence or absence of key public access computer resources and services in public libraries.

The information collected through the course of this study will be used by IMLS to help inform their policies and the support they provide to libraries. In addition, both public libraries and

public access computing centers will benefit from the information in this study since the results can point to different ways their policies and services can improve their assistance to the public. The collected information will also be relevant to schools offering programs in library and information science, where many library professionals obtain their educations. This study has implications for segments of the national public who have limited access to Internet and other information resources. Overall, the information collected for this study will help libraries and public access computing centers provide better service and information resources to the public.

A3. Use of Technology in Collection of Information

Data collection will take place in three stages: (1) a national telephone survey; (2) a web survey, and (3) case studies of 5 U.S. public libraries.

The national phone survey will be conducted by Telephone Contact, Inc. (TCI). The survey will employ a dual frame probability sample of households that combines a List Assisted random digit dialing (RDD) sample procedure with a Cell Phone exchange sample. The objective of the dual frame design is to increase the overall coverage of U.S. households in the survey because cell-phone-only households represent roughly 1 in 6 households in the U.S. in 2007¹. The overall goal is to complete 760 interviews with users of library public access computing with approximately:

- 600 of these respondents coming from the RDD sample frame,
- 80 respondents from the cell phone sample, and
- 80 from a non-response follow-up sample.

The RDD sample will also include an oversampling of telephone exchanges from low income areas in order to increase the number of interviews conducted with low income respondents.

Telephone interviewers will record survey responses via a Computer Assisted Telephone Interviewing (CATI) system that not only facilitates the administration of the screening and survey questions but streamlines the data collection process as well. The survey will require human interaction between respondents and interviewers.

The telephone survey will be an invaluable data source for developing estimates of usage rates among various important subpopulations (e.g., socio-demographic groups such as race, ethnicity, age groups, income levels, education levels, etc.). Generating such a portrait would be a major contribution to the library community and to policymakers. The primary difficulty with relying solely on the telephone survey for information about public access computer users is the low prevalence of users of public access computers in libraries² (8.9 percent of households used PAC in the past month in 2002) as well as the difficulty in gauging the public access computer resources available to respondents.

¹ In late 2007, an estimated 15.8% of U.S. households had cell-phone-only access as reported by Blumberg and Luke, 2008, "Wireless Substitution: Early release of estimates from the National Health Interview Survey, July – December 2007", NCHS, CDC.

² Glander, M. and Dam, T. (2006) Households' use of public and other types of libraries: 2002. NCES 2007-327, U.S. Department of Education, Washington, DC.

To assess the impact of resource availability on public access computer use and users and to generalize findings at the community and library levels, we will draw a random sample of public library systems proportionate to the size of the service population, oversampling library systems with lower than average levels of expenditures per capita as a proxy for low income communities. We expect a total sample of about 400 library systems will participate with the goal of completing a minimum of 200 online user interviews in each library system. The web survey will also enable us to supplement the telephone survey data by reaching low income or homeless persons who may not own telephones and a greater number of respondents in the 14-17 age range. We expect this to yield a total of 80,000 completed web interviews of PAC users in libraries.

Within the library systems, a random sample of public library administrative units will be drawn from the NCES database. Selected libraries will be asked to post a link to the web survey on their public access computer home pages and library websites. The Chief Officers of State Library Agencies (COSLA) association will assist in coordinating the web survey through state library agencies. The survey will be available at varying lengths of time depending on the target number of completes at each sampled library (due to our objective of oversampling users from libraries in low income areas as well as for our nonresponse follow-up).

The online survey instrument will be similar to the telephone survey, but will include fewer open-ended questions since participants are less likely to write in answers than they are to respond to verbal questions. The Social Development Research Group (SDRG) of the University of Washington will program and host the web survey. Data from the web survey will be encrypted and stored in a password protected SQL server database behind a secure firewall.

For the five case studies, interviews and focus groups will be conducted with library users, public library staff, and staff of peer agencies that might refer users to library PAC. Data will be recorded by the researchers using software on portable laptop computers.

The information requested in the telephone survey, web-based survey, and the case study data collection efforts has been held to an absolute minimum required to answer the research questions and minimize the burden on the respondents and cooperating libraries. When possible, items will be worded exactly the same on the phone survey and web survey to provide a basis for cross-validating individual items. All respondents will be asked to provide demographic data to control for factors such as income and education level and assess the degree of age, gender, and ethnic diversity in the sample.

A4. Duplication of Other Information

The information that will be collected for this study is unique, and therefore does not duplicate other efforts. Over the course of an extensive literature search, the researchers could not find previously collected information that explored the impact of public access computing in libraries.

A5. How Collection Impacts Small Entities

The web-based survey will impact small entities like libraries because qualified personnel like directors, librarians, and technical services personnel will be asked to link to the survey through their web sites and may be asked questions by survey participants. Libraries linking to the survey on their websites will be provided with several options for doing so that will minimize the effort required to accommodate this study. The survey has been specifically designed to be brief and to minimize the time burden of the survey and any possible burden on library staff. Further, skip patterns in the survey will allow respondents to whom the questions do not apply to answer fewer questions.

Case studies will involve analyzing the stakeholder landscape at 5 public libraries. COSLA will provide guidance in selecting case study sites that represent different geographic regions, demographic concentrations, levels of financial support for libraries, and best practices in public access computer administration. Participating state agencies will each receive \$2,000 to support their staff's time for preparation, promotion, and selection of libraries. Participating libraries will receive \$5,000 to support their staff's time for participation in interviews. Library users participating in case study interviews or focus groups will be given \$20.

A6. Impact of No Collection of Information

Once the model has been formulated, we will have a reliable basis for selecting the most appropriate indicators to use for measuring changes in the desired outcomes of public access computing at the individual, community and societal level. The final outcome will be a consistent and logical framework that can be used for both future research efforts in this area and for direct application to the desired policy outcomes of this study—providing data that can be used to support advocacy and funding for public access computers in public libraries.

The results from this study will have implications for public policy governing the provision of public access computer resources, especially for those who otherwise have limited or no access to electronic resources. Delays in collecting or not collecting this information at all would prevent researchers from understanding the ways in which public access computers are used by citizens and developing policies that are responsive to the needs of communities. The collection of this information is also warranted because the information this study proposes to collect is unique and will address holes in scholarly literature.

A7. Special Circumstances

There are no special circumstances that apply to this data collection.

A8. Federal Register Notice

A 60- day notice was published in the Federal Register December 2, 2007, vol. 72, no. 232, page 68199, and a 30-day notice was published in the Federal Register August 6, 2008, vol. 73, no. 152, page 457941 to solicit comments on the study to Assess the Impact of Access to Computers and the Internet and to Related Services at Public Libraries on Individuals, Families, and Communities prior to submission of this OMB clearance request.

A9. Payment/Gift to Respondents

Library users who participate in interviews or focus groups will receive a \$20 incentive in the form of an online gift certificate to Amazon.com for their participation. State agencies assisting with coordination with selected web survey sites will each receive \$2,000 to support their staff's time for preparation, promotion, and communication with libraries. Public libraries participating as case study sites will receive \$5,000 to support their staff's time for participation. Public access computer users completing the telephone survey will receive \$10 online gift certificates from Amazon.com upon completion of the survey. Respondents to the non-response follow-up study will receive \$2 enclosed with the notification letter and a \$40 online gift certificate to Amazon.com upon completion of the survey. Each library participating in the web survey will be provided a report analyzing the data collected from their users; those participating in the nonresponse study will receive \$200 each. No payments to web survey participants will be provided.

A10. Assurance of Confidentiality

Respondents to the telephone and web-based surveys and participants in focus groups and interviews will be advised that the reports prepared for this study will summarize findings and will not associate responses with a specific individual, and that identifiable information will not be provided to anyone outside the research team, except as required by law. They will be provided with information about the benefits of participation at the beginning of the survey or interview. The University of Washington Institutional Review Board has reviewed this data collection process and approved the instruments and methodologies being used.

The confidentiality of the data collected through the web survey is protected on many levels. Physical access to SDRG offices where servers are housed is controlled using entry codes to restrict access of non-employees. All computerized data is stored on secure servers with rights and permissions restricted and controlled through the central domain server. Computerized data is accessible only to authorized staff. After data is accumulated and verified, all identifiers on computerized data is destroyed. Additionally, a Certificate of Confidentiality has been previously received by this center from the National Institutes of Health (NIH).

A11. Justification of Sensitive Questions

While the goal of the national telephone and web surveys and the on-site interviews and focus groups is to identify, in very general terms, the type of information people access while using public access computers, there is a small subset of questions that ask respondents to report socio-demographic characteristics. Although this information may potentially be interpreted as sensitive, it is important to for the agency to gather this data to determine whether and how these social characteristics are correlated with certain types of public access computer use. Confidentiality assurances will be given to all respondents and data will be secured in accordance with accepted social science practice (see section A10).

A12. Hour Burden for Collection of Information

The estimated burden for the **telephone survey** is about *1,222 hours* and *\$29,609*³. This is based on an average 15-minute survey completion time for each of the 760 public library computer users and an average of 2-minutes to screen 34,400 households (based on 20,700 from the RDD, 3,200 from Cell Phone and 4,500 from the nonresponse follow-up) to find these individual respondents.

34,400 HH x 0.03 hours = 1032 hours
760 participants x 0.25 hours = 190 hours
1032 + 190 = 1222 hours x \$24.23 = \$29,609

The respondent burden estimate for the **web survey** is *20,000 hours* and *\$484,600*⁴. This is based on an average 15-minute survey completion time for each of an anticipated 80,000 respondents.

80,000 participants x 0.25 hours = 20,000 hours
20,000 x \$24.23 = \$484,600

The respondent burden estimate for the case study interviews is based on an expected respondent pool of approximately 250 library users, 50 staff of public libraries, and 50 additional stakeholders from local agencies that have referred the public to libraries for public access computer use. The estimated hour burden for **case study** interviews is *175 hours* based on a 30-minute interview time. The estimated cost burden for the interviews is *\$3,028* for library users, *\$613*⁵ for public library staff, and *\$482*⁶ for additional stakeholders.

250 users x 0.5 hours = 125 hours
125 hours x \$24.23 = \$3,028

50 library staff x 0.5 hours = 25 hours
25 hours x \$24.53 = \$613

50 stakeholders x 0.5 hours = 25 hours
25 hours x \$19.30 = \$482

A13. Total Annual Cost Burden for Collection of Information

There are no capital and start-up costs, or annual operation or maintenance or purchase of services costs to respondents.

A14. Annualized Cost to Federal Government

³ Based on 2006 Median household income (\$48,451) from <http://www.census.gov> with an hourly rate assuming 40 hours/week, 50 weeks/year.

⁴ Based on 2006 Median household income (\$48,451) from <http://www.census.gov> with an hourly rate assuming 40 hours/week, 50 weeks/year.

⁵ \$49,060 is the 2006 median annual earning of librarians from <http://www.bls.gov/oco/ocos068.htm>.

⁶ \$38,590 is the 2007 mean wages of community and social services specialists from <http://www.bls.gov/news.release/pdf/ocwage.pdf>

The estimated annualized cost to the federal government of this data collection effort is \$554,876.

A15. Program or Burden Changes

This is a new data collection.

A16. Plans for Tabulation and Publication of Collected Information

This data collection effort will commence immediately after OMB approval. We request that we receive approval for a 5-month data collection period from OMB. The timetable for key activities, demonstrating this need, follows:

Activity	Number of weeks after OMB approval
<i>Preparation for conducting the web-based survey</i>	<i>Weeks 1-4</i>
Develop sampling frame and select stratified pps sample of library systems	
Prepare/e-mail linking instructions to library administrators	
Prepare web-based survey system with approved questionnaire	
Conduct Pretest and revise instruments and field protocols	
 <i>Preparation for conducting the telephone survey</i>	 <i>Weeks 1-4</i>
Prepare survey system with approved questionnaires	
Select sample and partition into replicates	
Conduct Pretest and revise instruments and field protocols	
 <i>Field telephone survey</i>	 <i>Weeks 5-17</i>
Release sample in replicates	
Separate tracking of cell-phone versus RDD sample	
Non-response follow-up study	
 <i>Field web-based survey</i>	 <i>Weeks 5-17</i>
Field survey	
Monitor survey progress through weekly reports	
 <i>Process data files and develop frequencies</i>	 <i>Weeks 7-21</i>
Process survey data (ongoing, throughout data collection)	
Conduct nonresponse analyses (for weighting)	
Develop analytic weights	
Run series of cross-tabulations for analysis	
 <i>Select and prepare for case studies</i>	 <i>Weeks 1-4</i>
Draft memo recommending potential sites for in-depth study	
Finalize list of sites for in-depth study	

Conduct case studies

Weeks 5-18

Schedule and conduct site visits and focus groups

Write interview notes;

Process and analyze qualitative data

Develop and disseminate final report

Weeks 20-32

Draft outline for final report

Draft final report

Submit final report and briefing materials

Presentation of findings to IMLS

--

We anticipate the telephone survey will begin in the latter part of 2008, and take about two to three months to complete, with the case studies conducted concurrently. The web survey will take place over a two month period. Analysis of the survey questionnaire data will occur in the first quarter of 2009 while the analysis of the case study data will begin in the winter of 2008.

The report of the survey findings will include the survey methodology and the quality of the data, a description of the sampling procedures, and a discussion of problems encountered in administering the survey, and a calculation of standard errors and design effects for the key survey variables for all analytic domains. Principal quality and design parameters such as screening and eligibility rates for the telephone survey and cooperation and overall response rates for all surveys will be reported; the raw data counts used to calculate these parameters will be reported, as well.

It is anticipated that a final report will be submitted to the IMLS by the summer of 2009. Copies of the report will be provided on the IMLS and University of Washington Information School websites after final approval and release by IMLS.

A16.1 Analysis plan

In order to identify key areas of public access computing impact on individuals, families, and communities, a triangulated, mixed method analytic approach is warranted, involving both quantitative and qualitative analyses. Survey data will be used to produce statistically generalizable findings, principally in the form of tabular analyses for the overall sample as well as for specific substantively important domains such as age groups, race/ethnicity groups, sex, household income groups and geographic areas. The site visit interview and focus group data will complement the statistical survey findings and rely on the rich conceptual information to help understand analytic findings and stimulate policy relevant insights. Administrative and program data, such as publically available data on program support for public access computers by different government agencies and private foundations, will be obtained during site visits. This information will be used to provide context and clarification of data collected.

The three data collection methods will yield a large amount of different types of evidence:

- demographic data and responses to fixed-answer multiple choice questions that are easy to quantify,
- short responses to open-ended questions that are easy to code, and
- longer responses to in-depth interview questions and extensive field notes that will require iterative reading and multi-level coding.

Data analysis will comprise two phases: quantitative reporting of survey responses (in tabular form) and identification of themes in the qualitative analysis. The telephone survey data will provide a representative picture of the prevalence of different types of people using public access computers and how it benefits them. The web survey data, along with NCES data on the selected libraries' resources will allow for the analysis of the relationship between available public access computing resources and user results. The case study data will provide a richer understanding of how users use public access computing and the role it fills in their everyday information environment.

Stratified random sampling and offering user incentives for responding to the telephone survey will increase the likelihood of achieving a representative sample. The survey data will also be weighted to match key census demographic control totals (by using the collection of screened households regardless of eligibility to participate in the PAC user survey). After taking into consideration analytic weighting and survey design effects, we expect that the sample size for the telephone survey (n=21,000 screened households and 760 users) will result in a margin of error⁷ of about +/-4.5% with a confidence level of 95%. Similarly, after we account for clustering and weighting effects, we expect that the sample size of the web survey (n=80,000 respondents) will also result in a margin of error⁸ of about +/- 2.2% at the 95% confidence level and will allow us to generalize the results to other libraries. Note that the web survey data will offer substantial flexibility for subclass analyses, allowing an insightful portrait of PAC users to be developed.

For the telephone and web surveys, we will conduct data-screening to check for data-entry errors, inconsistencies, and identify missing cases for any systematic bias. The final samples will be tested for internal reliability, co-linearity, and intra-class correlation to assess reliability of the operational variables and validity of the findings. Initial analysis will consist of running descriptive statistics for all variables to identify the center and distribution within the population and bivariate statistics (correlation and cross-tabulation) will be used to test for associations between variables. For variables where the investigators identify a possible causal relationship based on the qualitative evidence, we will conduct path analysis (multiple regression) to determine the proportion of variance that can be explained by the relationship.

⁷ This represents the half-width of a 95% confidence interval of an estimated percentage near 50%, assuming one person per eligible household is selected, a weighting effect (i.e., Design effect due to differential weighting) of 1.5 and a nominal sample size of n=760; so that approximately, $4.5\% = 1.96 \times \text{Sqrt}[(1.5 \times 0.25)/760]$.

⁸ This represents the half-width of a 95% confidence interval of an estimated percentage near 50%, 200 interviews per library, 400 participating libraries (clusters), a conservative intraclass correlation of 0.15, a weighting effect (i.e., Design effect due to differential weighting) of 1.25 and a nominal sample size of n=80,000. Using $\text{DEFF} = (1 + [b-1]\text{roh}) \times 1.25$, where roh is the intraclass correlation, this yields $\text{DEFF}=38.6$ so that approximately, $2.2\% = 1.96 \times \text{Sqrt}[(38.6 \times 0.25)/80,000]$.

For the qualitative data, we will analyze data as they are collected, following an approach that will aid in identifying a range of responses for each indicator (Lofland & Lofland, 1995; Miles & Huberman, 1994). The schemes will reflect the data's emergent themes and will be guided by the study's logic model. A code book will be used to assign terms to all segments in the data that reflect particular concepts. After the final schemes are developed, tests of inter-coder reliability will be conducted with independent coders and final adjustments will be made to the codes (c.f., Krippendorff, 1980).

To ensure trustworthiness (reliability and validity) of the qualitative data, we will use several measures (c.f., Chatman, 1992; Lincoln & Guba, 1985). Reliability will be ensured through: (1) consistent note-taking, (2) exposure to multiple and different situations using triangulated methods, (3) comparing emerging themes with findings from related studies, (4) employing intracoder and intercoder checks, and (5) analyzing the data for incidents of observer effect. Validity will be assessed as follows:

- Face validity: ask whether observations fit an expected or plausible frame of reference;
- Criterion/internal validity (credibility) based on pre-testing instruments, rigorous note-taking, methods, peer debriefing, and member checks or participant verification;
- External validity: provide "thick description" and comprehensive description of our methods so others can determine if our findings can be compared with theirs;
- Construct validity: examine data with respect to public access computing outcome literature, models of public library use, and principles of information behavior

The research procedures have been reviewed by the UW Internal Review Board and comply with federal regulations regarding the protection of human subjects participating in academic research. Subjects will be informed of their rights as participants, and may refuse to answer any question or end their participation at any time. All responses will be anonymous; no identifying information will be recorded. Subjects will be at minimal or no risk of suffering stress, embarrassment or discomfort from this study. Children under age of 14 are not competent to give legal assent, thus they are ineligible to participate.

A16.2 Data presentation

The preliminary report will include an executive summary, literature review, statement of methodology, analysis, report on findings, and recommendations (including descriptions of how the results can be used by practitioners to measure the impact of their public access computer services, improve the services offered and to understand how the complex array of decisions they make about public access computer work as a whole to serve individual users and the community). The final report will include complete technical information: instruments, full data summaries, and detailed description of methodologies used. The synthesis will summarize the background and findings in a national context to be informative to a broad policy and planning audience. The results will also be disseminated through professional and academic conferences and journals.

A17. Expiration Date

The OMB approval number and expiration date will be displayed on all survey instruments and discussion guides.

A18. Certification Statement

There are no exceptions to the certification.

References

Chatman, E. A. (1992). *The information world of retired women*. New York: Greenwood Press.

Chute, A. and Kroe, P.E. (2007). *Public Libraries in the United States: Fiscal Year 2005 (NCES 2008-301)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

IMLS. (2007). Program solicitation for a cooperative agreement to assess the impact of free access to computers and the internet at public libraries (Funding Opportunity No. GF-CA-07). Available at <http://www.ims.gov/pdf/ComputerAccessStudyRFP.pdf>

Jaeger, P. T., Bertot, J. C., & McClure, C. (2007). Public libraries and the Internet 2006: Issues, findings, and challenges. *Public Libraries*, 46 (5), 71-78.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Newbury Park, CA: Sage

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Lofland, J., & Lofland, L. H. (1995). *Analyzing social settings: A guide to qualitative observation and analysis*. Belmont, CA: Wadsworth.

Manfreda, et al. (2008). "Web surveys versus other survey modes: A meta-analysis comparing response rates." *International Journal of Market Research*, 50(1).

Miles, M. B., & Huberman, A. M. (1994). *Qualitative analysis*. Thousand Oaks, CA: Sage.

Zweizig, D., & Dervin, B. (1977). Public library use, users, uses: Advances in knowledge of the characteristics and needs of the adult clientele of American public libraries. *Advances in Librarianship*, 7, 231-255.