上海图书馆
SHANGHAI LIBRARY
Add:1555 Huai Hai Zhong Lu
Shanghai 200031 China
Tel:021-6445 5004 Fax:021-6445 5006

Arthur Hendricks
Assistant Systems Librarian
Assistant Professor
Branford P. Millar Library
Portland State University
951 SW Hall, Post Office Box 1151
Portland, Oregon 97207-1151
U.S.A.

March 21, 2002

Dear Mr. Arthur Hendricks:
Libraries have entered a new age of development. Currently, knowledge navigation in library services is not only an eye-opener or research topic but also an emerging practice in libraries' day-to-day operation.

For exchanging of the views and experiences, if any, on the subject, as well as for commemorating the 50th anniversary of the Shanghai Library, the Shanghai International Library Forum will be held from July 15 to 18, 2002. The theme of the Forum is "Knowledge Navigation and Library Services: International Cooperation".

It is my pleasure to cordially invite you to attend the Forum and make a presentation on the Forum. We shall be very pleased if you could send your topic and abstract by March 30th, 2002 and the paper by April 15th, 2002. The proceedings are to be published before the Forum.

While the Organizing Committee is not in a position to provide financial support for you to attend the Forum, we would be pleased to facilitate your attendance by assisting with hotel reservations.

I'm looking forward to meeting you in Shanghai.

Dr. Wu Jianzhong
Director
Shanghai Library
P. R. China

以因特网为研究工具，增进图书馆员间的信息共享

Jian Wang
（美国波特兰州立大学图书馆连续出版物编目馆员、助理教授）
Arthur Hendricks
（美国波特兰州立大学 Branford P. Millar
图书馆系统助理馆员、助理教授）

摘要：因特网的产生和发展为学术研究提供了许多新机会。越来越多的图书馆员也以因特网为工具来收集数据，探讨课题，以及交流信息。此文将讨论一例利用万维网作为调查工具对当今图书馆为计算机数据存储器的选择和使用的调查。此研究的目的在于发现各类图书馆（包括大学，公共，公司，中小学，政府机构，以及法律部门图书馆）当前对计算机数据存储器的选择以及今后在这方面的打算和计划。本调查的对象是世界各地的图书馆员。接受调查问卷的图书馆员来自美国，加拿大，英国，澳大利亚，香港，台湾，以及新加坡等国家和地区。本文作者详细讨论了如何利用因特网进行问卷调查，收集和分析数据，以及对此调查结果的理解和分析。本文也对利用因特网作为调查工具来收集数据的优越性和问题作了一些讨论。

Using the Internet as a Research Tool for Information Sharing Among Librarians Across the World

Jian Wang
(Serials Catalog Librarian
Serials/Documents Cataloging Coordinator
Assistant Professor
Portland State University Library, USA)
Arthur Hendricks
(Assistant Systems Librarian
Assistant Professor
Branford P. Millar Library, Portland State University, USA)

Abstract: The emergence of the Internet offers new opportunities for research endeavors. Increasingly, the Internet is being used by librarians as a research tool to gather data, investigate issues, and share information. This paper focuses on the utilization of Web-based surveys for conducting library research. Discussions are based on the authors' recent research regarding libraries and desktop storage options. The survey research aimed at finding current and future storage options at all types of libraries (academic, public, corporate, school, government, and law).

Populations of the study are librarians across geographic borders, including the U. S. , Canada, England, Australia, Hong Kong, Taiwan, and Singapore. The authors describe how the Internet can be used as a research tool to conduct the survey, analyze the collected data, and present the survey results. Practical issues in designing and implementing Web-based surveys are highlighted. Technical details are illustrated. The advantages of using the Internet for obtaining research data, as well as concerns associated with the use of Web methodology are also included.

1 INTRODUCTION

One of the top technology trends for libraries in the 21st century listed by the American Library Association, Library and Information Technology Association (LITA), is that "libraries need to work with the Internet" and librarians need to make it a top priority in learning about the Web for innovative ideas and applications (http://www.lita.org/committee/toptech/mw2000.htm). Indeed, librarians were one of the primary professional groups to embrace the new technologies of the Internet. Digital libraries, electronic databases, and virtual reference are some of the examples. Conversely, libraries have benefited a great deal from the Internet as a tool in providing access to online-information services, delivering instruction for end-users, and communicating with other professionals, libraries, and the community.

This paper focuses on the innovative use of the Internet as a research tool to obtain data, investigate issues, and share information among library professionals across the world.

Discussions are based on the authors' recent research regarding libraries and desktop storage options. The survey research aimed at finding current and future storage options at all types of libraries (academic, public, corporate, school, government, and law).

Populations of the study are librarians across geographic borders, including the U. S. , Canada, England, Australia, Hong Kong, Taiwan, and Singapore. The authors describe how the Internet can be used as a research tool to conduct the survey, analyze the collected data, and present the survey results. Practical issues in designing and implementing Web-based surveys are highlighted and technical details are illustrated. The advantages of using the Internet for obtaining research data, as well as concerns associated with the use of Web methodology are also included.

2 WEB-BASED SURVEYS VS. PAPER SURVEYS

Traditionally, most researchers utilize a mail questionnaire to gather data in survey research. This survey method, also called paper surveys, has many disadvantages. It is time consuming and costly. Paper surveys have to be reproduced (either printed multiple times or photocopied), stuffed into envelopes, stamped, and mailed out.

Return postage must be included to aid in getting the surveys returned and yet the response rate tends to be lower than electronic surveys. It is also labor-intensive for both the researcher and the subjects. In addition, chances for errors entering data from returned surveys are likely to happen.

The Internet offers new possibilities for research endeavors. Utilizing the Internet as a research tool, such as Web-based surveys, to obtain data critical to performing research is becoming increasingly popular among librarians. The Web-based surveys offer many unique advantages than paper surveys do. They include: inexpensive distribution (no postage costs); time saving; quick delivery (one email goes to many electronic email lists); no barriers of geographic distance; easy for subjects to respond by a clickable link; responses are automatically entered in a machine-readable form; no chances for data entry errors in the traditional manner; and easy to import data from flat ASCII text into an Excel spreadsheet.

In addition, with regard to the validity of the Web-based studies, researchers found that both the Web-based surveys and the paper surveys achieved similar responses and/or similar internal consistency (Perkins and Yuan, 2001).

3 LIBRARIES AND DESKTOP STORAGE OPTIONS SURVEY

In November 2000, the authors conducted a Web-based survey to investigate current and future desktop storage options in libraries with the hope of helping librarians in deciding on the best solutions for their institution, given the fact that technology is fast changing and floppy disk drive may soon become obsolete. The survey was developed using a single, static html page and sent to nine electronic discussion lists representing

a random and representative statistical sample to librarians in all types of libraries without regard to geographic location. A total of 248 responses was received and the results revealed that libraries, especially academic libraries, do not have a current plan to cope with the forecast obsolescence of the floppy disk, but most libraries are willing to install floppy disks if PC manufacturers stop installing them. Our survey also found that if the floppy disk is discontinued, most libraries would want to have such choices as (1) the Local Area Network (LAN), (2) Zip disks, (3) CD-RW or (4) hard drives for replacement.

Designing and implementing a survey via the Web is a challenging and complicated process because it require more than research skills. Knowledge of HTML, PERL, and CGI programming is necessary to be able to construct Web-based survey forms and manipulate collected data for analysis. Described below are the steps involved in designing and implementing our recent Web-based survey.

3.1 Identify the focus of the study and method of research

Our first step is to decide what we want to know from the study, so we can define research goals and objectives. We decided to investigate how libraries currently use desktop storage options for staff and patrons and if they have any plan for future options, as the computing needs for data storage have been increasing and the traditional floppy disk may become obsolete soon. A Web-based survey was the chosen method for this study.

3.2 Develop survey questionnaires

The next step is to design questionnaires. Our experience suggested that it is important to know what answers you are looking for when constructing a survey questionnaire. We found the following guidelines (Janes, 1999) helpful in writing a good, unbiased and answerable question.

(1) Write questions that are only related to the problem at hand.

(2) Use the correct type to get the best information, such as multiple-choice, open-ended or closed-ended.

(3) Avoid jargon and give clear, unambiguous, precise definitions.

(4) Do not lead the subjects by saying "Don't you think that..."

(5) Make sure that questions can be answered by the subjects.

(6) Ask one question at a time, avoiding the use of "and" for two separate questions.

(7) Keep questions short.

(8) Avoid the word "not" in a question, it can easily be misread and thus changes the question.

(9) Do not impose biased influence.

Examples from the survey:

While at work, what do you currently use for desktop storage? (Check all that apply) (multiple responses)

- Floppy disk ☐
- Tape drive ☐
- Hard drive ☐
- Zip disk ☐
- Superdisk (LS-120) ☐
- CD-RW ☐
- DVD-RW ☐
- Network (Local Area Network) ☐
- Internet drive (e.g. Xdrive) ☐

- RAID ☐
- Other (please specify) ☐
- Don't know ☐

On a scale of 1-5, 1 being not important, please rank the impact of the forecast obsolescence of the floppy disk: (scale responses)

☐ 1 not important

☐ 2

☐ 3

☐ 4

☐ 5 very important

3.3 Develop html files (static vs. dynamic)

Once the questions have been written, the next step is to design the instrument. In this study, a Web-based form was developed. Although there are a number of commercial programs available that help construct survey forms on the Web, the authors decided to create their own form to minimize the cost and to exercise expertise as well. A single, static html page was created with a total of 15 questions (plus a place for additional comments). A Common Gateway Interface (CGI) script written in PERL 5 parsed the form data. The script sent a response in plain English to the authors via email as well as appended a file that was machine-readable. The resulting file saved the responses in a flat ASCII file that was later imported into Microsoft Excel.

As mentioned above, a static page is one that is created and does not change. It looks the same regardless of the answers given in the survey. A dynamic page, on the other hand, uses a scripting language and reacts to the answers given. In the authors' survey, the page was static and simply recorded whatever the respondent typed. Perhaps a better approach is to create a script that changes the questions depending on the answers. For example, we asked if the respondent had any plans for the obsolescence of the floppy. If yes, we asked what those plans were. Even if the person said they had no plans, they still read the question asking what the plans were. If we had used dynamic html, as soon as the person said he/she had no plans, the page would skip over the question asking for details and move onto the next subject.

3.4 Pretest the survey instrument for content and design

Once the bugs were worked out of the script, several test subjects were given the survey as a pretest. These people took the survey and then proffered feedback to the authors, who in turn made changes to the survey and the CGI script. The purpose of the pretest is to insure the effectiveness of the questionnaire and usability of the Web form.

Obtain approval

The authors contacted human subject studies for permission of this study. One should always be aware of the rules regarding surveys. If you are conducting a survey and the subjects are minors, you need to get approval. One usually has to submit the survey and describe how the data will be used to get permission to publish results. If you simply want to improve a service and do not plan on publishing the results, you usually do not have to seek permission. It is best to find out the rules on your campus first before you begin surveying.

3.5 Select the survey sample

Once the final version of the survey was complete, the survey population was defined. The authors identified nine technology-related electronic discussion lists for soliciting responses to the survey. They include WEB4LIB (for library-based World Wide Web managers), libs-or (Oregon libraries), ASIS&T (American Society for Information Science and Technology), CALA (the Chinese American Librarians Association), INNOPAC list (for INNOPAC users), LIL (Linux in Libraries), siglan (special interest group, local area

networks, ASIS&T), and SLA-DITE (the Information Technology Division of the Special Libraries Association). These groups were selected because they were representative of the population that we wished to survey. Some of the lists are international in scope, so the sampling population is worldwide. The authors received responses from such countries as Canada, England, Australia, Hong Kong, Taiwan, and Singapore, although the majority of the respondents were from the United States.

3.6 Implement the survey instrument

An email message was sent out to the identified electronic discussion lists with a link to the survey form. The email serves as a cover letter with instructions to the survey and a confidentiality statement. The survey was sent out several times as a reminder and the subjects were asked to complete and return the survey by a given date. The time given for completing the survey was three weeks for this study. It is important to stress a deadline, otherwise people will put off responding to your survey and you will not get many responses.

Example of the email message:

"Technology is changing fast. It is predicted that the 3.5-inch diskette drive on PCs will soon be discontinued by the computer industry. PCs with CD/DVD drives are becoming standard. Libraries are challenged to rethink their desktop storage options in preparation for this change. The purpose of this survey is to gather information on the desktop storage options used or to be used in libraries.

The survey is at http://www.lib.pdx.edu/bvah/survey/survey.html and the deadline is April 1, 2001. Responses to this survey will be summarized, and the results will be sent to those who participate in this survey if requested. All answers will be kept confidential. This survey should take only 5 minutes to complete.

Thank you very much for your interest and time in participating in this survey."

3.7 Analyze the data and interpret the results

Survey responses were collected and entered directly into MS Excel files via a flat ASCII file. Data were then analyzed using SPSS statistical software once the data was converted into values that SPSS could handle. For example, the script recorded "systems librarian" which SPSS could not use. We had to assign an integer to job titles in order for SPSS to do cross tabulation. Cross-tabulated tables were used to examine the relationship between two or more variables. It is interesting to explore the influence one variable may have on another.

The t-test determines whether the differences between group averages could have occurred by chance. A .05 level of significance means that there are 5 chances of one hundred trials that a difference would occur by chance. One precaution that should always be considered when using a t-Test: the sample data should be drawn from populations with normal distributions. Fortunately, this is often the case, especially with physical property measurements of materials.

For example, we used a t-test to determine significant differences in the averages between academic and non-academic libraries and found that CD-RW, Network and Internet drives were picked most by academic libraries versus all others as potential replacements for the floppy disk. Since a value of one, the lowest number, indicates the most preferred choice, lower averages indicate preferred choices. Non-academic libraries ($M = 2.41$)① rated CD-RW lower than academic libraries ($M = 3.11$). Academic libraries ($M = 2.56$) rated network storage more than all other libraries ($M = 2.03$). However, academic libraries rated Internet drives ($M = 3.14$) less than everyone else ($M = 4.71$). These differences are all statistically significant ($p < 0.05$).

4 TECHNICAL CONSIDERATIONS

In this paper, we assume that you have an understanding of the World Wide Web and are familiar with programming. Creating the web pages with your questions for a survey is not that hard if you use a web editor

① Big M's are the Mean.

such as Microsoft's FrontPage or Macromedia's Dreamweaver. You can also simply use Netscape's Composer to create the page. It is more important to focus on the actual survey instrument and not the technology behind it.

Programming on the web involves using a Common Gateway Interface (CGI) script written in a programming language such as PERL. PERL stands for Practical Extraction and Report Language. It was designed by Larry Wall as a program tool and is similar in syntax to the C programming language. It is also like a shell script language. It runs on UNIX as well as other operating systems such as Windows and Novell Netware. If you have access to an account on a UNIX server, chances are it already has PERL. If you have access to a command prompt, type which PERL or where PERL and it will give you a path if it finds it such as /usr/local/bin/perl. There are many pre-made PERL scripts available for free download on the web. PERL scripts are easily modified with a text editor. You can use vi or emacs or even pico to edit the scripts and then run them. PERL is compiled every time you run the program (it is interpretive not complied).

You can easily test your script and make changes to it since it is an interpretive language. If something doesn't work, simply open the script in your text editor and fix it and try the script again. While connected to the server, you can test the syntax of your script simply by typing perl -c name-of-your-script. pl to see if there are any errors before actually running your program.

The web page can be static or dynamic. For our floppy research survey, we used a static HTML page that was linked to a PERL script. It is linked through the <form> tag. There is a start and close to the tag, and between the opening tag are the two tag attributes, method= and action=. The method= attribute configures the web browser to send the form's information to the Web server. The action= attribute tells the browser where the form's matching PERL script is located on the Web server, so that the script can process and react to the information submitted via the HTML form.

All forms are made up of elements and most form elements have two attributes name= and type=. Form elements identify the form's different types of information and methods for delivering information. Each form element must have a unique name so that the server will be able to identify each entry as a separate variable.

The CGI script then records what each variable is when the user enters his or her data into your form.

```
# Get the input
    if ( $ENV{'REQUEST_METHOD'} eq 'get') { $buffer = $ENV{'QUERY_STRING'}; }
    else { read(STDIN, $buffer, $ENV{'CONTENT_LENGTH'}); }
    @nvpairs = split(/&/, $buffer);
    foreach $pair (@nvpairs)
    {
    ( $name, $value) = split(/=/, $pair);
    $value = tr/+/ /;
    $value = s/%([a-fA-F0-9][a-fA-F0-9])/pack('C', hex( $1))/eg;
    # Stop people from using subshells to execute commands
    $value = s/! / ! /g;
    # Uncomment for debugging purposes
    # print 'Setting $name to $value<P>';
    $FORM{ $name} = $value;}
```

This part of the script extracts the data submitted from the HTML form and then assigns the values to variables. So when we ask "What type of library do you work in," <input type=radio value=academic name =library>, and someone selects the radio button for academic, there is a variable called "library" with the value of "academic." If you are going to use SPSS, it would behoove you to assign integers for the values instead of phrases as this is what SPSS uses.

Once all the fields are parsed earlier in the script, another section writes the files out to a delimitated flat

ASCII file. The script is very simple:

```
open (ANSWERS, '>> $full_path') || die ('Couldn't open $full_path');
flock (ANSWERS, $exclusive_lock);
print ANSWERS <<End_Of_Write;
$print_count ^ $FORM{'library'} ^ $FORM{'other1'}^ $FORM{'position'}^[all of the variables go in here separated by carets (^)]End_Of_Write
flock (ANSWERS, $unlock);
close(ANSWERS);
```

The script here just prints whatever is between the End_Of_Write tags. The caret (^) is used so that when the ASCII file is imported into Microsoft Excel, the data stays in its own field. Other symbols could be used, but since we were collecting people's responses, we had to use something that a person would likely not use, such as a comma.

The script then sent an email message in a more human-readable form.

```
# format the mail file
format MAIL =

^<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
$value
. # Open the mail file and write to it
open (MAIL, '| $mailprog $recipient') || die '$mailprog not available.\n';
print MAIL 'Reply-to: $FORM{'email'} ($FORM{'name'})\n';
print MAIL 'From: $FORM{'email'}\n';
print MAIL 'Cc: \n';
print MAIL 'Bcc: $me\n';
print MAIL 'Subject: Floppy Survey Response $print_count\n\n';
print MAIL '\n';
# type of library start 1
if ($FORM{'library'} = /[^\r\t\n\f]/)
{
print MAIL 'Library: $FORM{'library'}\n\n';
}
if ($FORM{'other1'} = /[^\r\t\n\f]/)
{
print MAIL 'Library: $FORM{'other1'}\n\n';
}
# type of library end
```

In the beginning of the script, we define which mail program to use to send the mail. Since we were using a UNIX account, we chose sendmail. The $ print_count variable comes from a separate file that is written to and which is used to track the responses. This wasn't necessary, but it was useful when receiving email to see how many responses we were getting. The various fields in an email message have to filled in either from the form (such as name) or generated by the script (such as the email address — these are supposed to offer anonymity, so you can have the sender's email address be yours). The if/then statements are used to print only the variables that have data. Since the email responses are supposed to be human-readable, there is no need to send blanks to the recipient.

Data goes into a machine-readable ASCII file and in human-readable email (depending on the script), im-

port ASCII into MS Excel using its import feature. ^ (caret symbol) as deliminator (avoids problems with text answers that might contain punctuation).

An example of the data generated follows:

00248^ academic^ ^ ^ reference librarian^ floppy disk^ ^ hard drive^ don't know^ not applicable^ not applicable^ not applicable^ not applicable^ not applicable^ not applicable^ not applicable^ not applicable^ not applicable^ ^ not applicable^ don't know^ ^ ^ ^ ^ ^ don't know^ no^ ^ yes^ ^ yes^ not applicable^ 5^ 1^ 4^ 2^ 5^ 5^ ^ ^ ^

Imported into MS Excel, it looks like this:

248	academ ic			reference librarian	flop py disk		hard drive								

					don't know	not applicable	not applicable	not applicable	not applicable	not applicable	not applicable

not applicable	not applicable	not applicable		not applicable	don't know						don't know	no	

yes		yes	not applicable	5	1	4	2	5	5			

IN RETROSPECT

While the Web offers great potential for collecting and analyzing data, it also has created challenges and presented some drawbacks. It requires Web access, a server, and knowledge to set up a script (some UNIX accounts don't let you into the shell). Most people do not have access to the cgi-bin folder on their university's web server, so they must put the script in their personal directory. This is assuming that you are allowed to. If you are, then it is better to use . cgi as the file extension for the script versus . pl. The problems, such as multiple responses from one person, returned entries from unknown population sample, blank responses were not caught in the script due to technical difficulty, need to be addressed.

If we conducted another Web-based survey, we also would consider performing an error checking so that the respondents were forced to answer questions. In some instances, respondents were given a chance to fill in blanks if they made a particular choice that needs explanation. Some would indicate a special circumstance but fail to express what that was. Error checking would prevent non-responses and is easily accomplished in PERL. PHP is another scripting language that could be used to automate this process. PHP can be faster to produce than PERL Also, PHP can be used in the same directory as the user's HTML files.

In retrospect, we would definitely use the Web again to conduct future surveys. There is no cost as in mailing questionnaires and no chance of entering data incorrectly. It is very easy to import plain text into Microsoft Excel and import the resulting spreadsheet into SPSS. However, it is important to remember that, unlike mailed surveys, there is no way of knowing the return rate. One could make an educated guess as to how many people are subscribed to a particular discussion list and assume that the email addresses are valid and that of those are the recipients see ones message advertising the survey.

One could also use a dynamically generated survey depending on the answers. For instance, if you were surveying academic libraries to see what kind of electronic reference they offer, the form could stop asking questions if the respondent indicates that they do not offer it. Otherwise, the whole page will ask all sorts of questions that do not apply.

5 CONCLUSION

As the Internet continues to expand and shape our future in the information age, the library community is

becoming increasingly global. International collaboration and cooperation is becoming more and more important in today's connected and networked world. An example of this trend is the use of the Internet as a communication tool to share and exchange information among librarians worldwide. Our Web-based survey has demonstrated the benefits of utilizing the Internet technology to conduct research from a global perspective and be able to share information with other professionals without geographical barriers.

REFERENCES

Gay Helen Perkins and Haiwang Yuan, "A Comparison of Web-based and Paper-and-Pencil Library Satisfaction Survey Results," *College & Research Libraries* 62/4-3 (July 2001): 1.

Joseph Janes, "Survey Construction," *Library Hi Tech*, 17/3 (1999): 2—3.

Louis M. Rea and Richard A. Parker, *Designing and Conducting Survey Research: a Comprehensive Guide*, 2nded., (San Francisco: Jossey-Bass Publishers, 1997): 5, 12, 45—52.

Alsehli, Fafad, "The Use of the Internet as a Research Tool by Faculty at the Schools of Information Studies (IS) in the United States." Ph. D. diss., Florida State University, 1999.

Junion-Metz, Gail, "Instant Web Forms and Surveys for Academic Libraries." New York: Neal-Schuman, 2001.

Powell, Ronald, "Basic Research Methods for Librarians." 2nd ed., New Jersey, Ablex Publishing Corporation, 1994.

Web survey methodology <http://www.websm.org/topics.html> [April 21, 2002]