Surveying the Territory: GVU's Five WWW User Surveys

Colleen M. Kehoe
Graphics, Visualization, & Usability Center
Georgia Institute of Technology
Atlanta, GA 30332-0280

James E. Pitkow
Graphics, Visualization, & Usability Center
Georgia Institute of Technology
Atlanta, GA 30332-0280

Citation: Colleen M. Kehoe & James E. Pitkow, Surveying the Territory: GVU's Five WWW User Surveys, *The World Wide Web Journal*, Vol. 1, no. 3

WWW User Survey Home Page | For more information send e-mail to www-survey@cc.gatech.edu

---

Introduction

Five years is not very long on most historical scales, but for the World Wide Web (WWW) it constitutes a lifetime. A question almost as old as the web itself is, "Who is using it, and for what?" One way to answer this question is to use paper surveys, telephone surveys, or diaries which are some of the the same methods used to measure the audiences of other one-way media such as television and radio. However, something interesting happened in early 1994: the implementation of HTML Forms turned the web into a two-way medium which made it possible to contact the audience directly. To test the viability of the web as a survey medium and collect preliminary data on the web population, the first GVU WWW User Survey was conducted in January 1994. Subsequent surveys have been conducted approximately every six months. The collection of responses from over 55,000 Web users over five surveys has given us a unique perspective on the advances in surveying technology and methodology and changes in the web population itself. In the following sections, we discuss what we have learned in each of these areas.

Evolving Technology

With each survey, we have attempted to advance the state of surveying technology and take advantage of new Web capabilities. Our first survey was the first publicly accessible Web-based survey and it pushed browsers that supported HTML Forms to their limit. Error reports from respondents and garbled results in our database quickly revealed differences between the various browsers in their handling of Forms. Although most major differences have been resolved, some minor ones persist and as a result, the data usually contains some errors which much be corrected by hand.

Many important features of the architecture were introduced in the second survey: adaptive questioning, enforcing questionnaire completion, and allowing user-selected IDs. Adaptive questioning means that the questions a user is asked depend on his or her answers to previous questions. 1 Questions adapted in a "batch mode" using CGI scripts: respondents answered a set of questions, submitted their answers, and got back a new set of questions which were follow-ups to the ones they had submitted. Also, questions from the original set which the respondent did not answer were returned along with the follow-up questions. Questionnaires were not accepted until all of the questions asked had been answered, preventing users from accidentally skipping questions. (For sensitive questions, we provided a "Rather Not Say" option.) The final feature introduced in the second survey was user-selected IDs which were used to relate a particular user's answers across different sections of the survey. After entering an ID, users were given a URL which contained their ID to add to their hotlist and which they could use to participate in future GVU surveys. This simple mechanism would allow us to do a longitudinal analysis of users who participated in several surveys.

Longitudinal tracking was tested during the third survey and fully deployed during the fourth. When users identified themselves as having previously participated in a GVU survey, either by using the URL they had stored or by remembering their ID, we used a weak challenge-response mechanism to verify their identity. Users were asked their location and age during the last survey and if their responses matched those in our database, we considered them verified. Note that this was not an attempt at true, reliable authentication; it was simply designed to minimize errors in identification and to discourage blatant mis-identification attempts. Users who did not want to participate in the longitudinal study were asked to simply choose a new ID for each survey. To make it more convenient for users who participate in multiple surveys, we filled in as much of the general demographics questionnaire as possible with their previous answers. Users could then simply review their answers and change them when necessary.

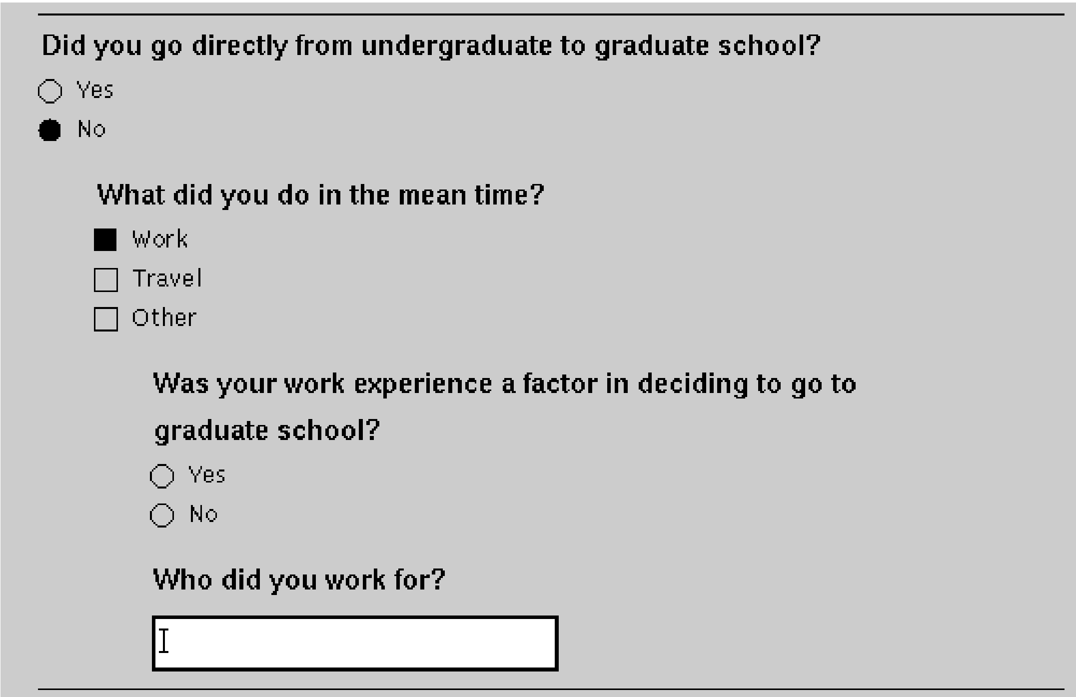
The recent introduction of Java to the web has opened a variety of possibilities for improving the survey technology. Originally, because of the web's limited interactivity, the vision of truly adaptive questioning could not be fully realized. While "batch mode" adaptability was a reasonable solution, it did not have the natural, conversational progression of questions we were aiming for. We felt a natural progression of questions would help respondents to give better answers because the proceeding questions would provide a context for the current question. Java made it possible to have the desired degree of adaptability--each mouse click had the potential to trigger new questions which could be asked immediately. To test this idea, we implemented a prototype survey applet which was offered as an option in the fifth survey. Since the survey applet has not been discussed in any of our previous publications, we describe it in more detail in the next section.

Prototype Survey Applet

There are three distinct portions of the survey applet: the adaption engine, the user interface, and the server interface. At the heart of the adaption engine is a simple production rule system. The survey designer specifies the way the survey adapts by creating a set of rules of the form: "if the answer to question X is A, then ask question Y". Rules may have multiple conditions, "(X is A) and (Y is B) and (Z is C)" but they can only test for equality and not arbitrary expressions. Every time a user answers a question, a "fact" is asserted, such as "the answer to question Z is D". The list of rules is then evaluated to see if this new fact satisfies any of the conditions. When all of the conditions for a rule are satisfied, the rule "fires" and the new question is added to the list of questions currently asked. If the question is already currently asked, it is not asked again. This situation can occur if the same question appears on the right-hand side of more than one rule. Questions which have no conditions for being asked (i.e. the initial questions) are given a condition of "NIL" which is always considered to be satisfied. Facts may also be retracted if a user changes an answer to a question. Any questions which were asked as a result of the fact having been asserted are then "unasked".

When the applet is loaded, the questions and adaption rules are read in and the initial questions are displayed. The applet supports the standard types of survey questions: checkbox, radio button, scrolling list, selection pop-up, and text entry. All of these can have follow-up questions which, when triggered, are placed on the screen slightly indented and immediately following the question that triggered them. Another strategy would be to append the new question to the end of the survey. We chose the first placement strategy because it places the new question (or questions) in the user's current area of attention, making the connection between the user's action (clicking) and the system's action (adding the question) explicit. Connecting a particular answer to a particular follow-up question also helps the user understand why that question is being asked and provides a context for interpreting it. As with the Forms version of the survey, the applet enforces question completion. If a user tries to submit an incomplete questionnaire, the unanswered questions are highlighted in red and can be easily spotted when scrolling back through the survey.

The applet integrates seamlessly with the CGI scripts used to collect the results. To submit the results, the applet creates a URL which mimics the format of Forms output. The name-value pairs are created from the answers to the currently asked questions and appended to the URL for the CGI script to create a GET-style URL. The applet then calls the *showDocument()* method on this new URL, submitting the results and returning the user to the same point they would have reached if they had used the Forms version of the same questionnaire.



Evaluation of Prototype

Making the survey applet available to a large number of users revealed some interesting technical issues. First, we expected that use of the applet would decrease the load on our web server since adapting questions and verifying completeness could be done locally, without a call to the server. This was not the case, however, because each "module" needed by the survey applet was retrieved by a separate call to the server. In object-oriented programming, programs consist of a set of objects. In Java, these objects are called "classes" and each class is stored in a separate file. There are approximately 25 classes used in the applet which currently must be retrieved with 25 separate calls. Not only did this make loading the applet very slow, but it required many more calls to the server than the Forms version does. (Granted, the classes only have to be retrieved once even if the applet is run several times for different questionnaires. Even taking this into account, the Forms version requires fewer calls to the server to complete the main part of the survey.) Having fewer, but larger and more complex classes is one option, but it violates good object-oriented design guidelines. A better solution to this problem would be to allow an entire set of classes, perhaps in a compressed format, to be retrieved with only one call to the server. Another interesting revelation, which is probably well-known to most Java programmers by this point, is that different browsers (and different versions of the same browser) do not handle Java applets the same way thereby defeating their cross-platform benefits. This situation is very similar to the differences in the handling of Forms a few years ago and will probably be resolved as these browsers become more mature.

Evolving Methodology

How can we ensure that the results of the survey are meaningful and valid? The field of distributed, electronic surveying is still very new and consequently any results obtained must be interpreted conservatively. Our survey suffers two problems that limit our ability to generalize from the results: self-selection and sampling. When a person decides to participate in a survey, they select themselves. There is very little researchers can do to persuade someone to participate if they simply prefer not to. The potential problem is that this decision not to participate may reflect some systematic judgment by a segment of the population being studied, causing them to be excluded from the results. However, all surveys have this problem to some extent; when a potential respondent hangs-up on a telephone survey or does not return a direct-mail survey, self-selection has occurred.

The more fundamental problem is sampling. There are basically two types of sampling: random and non-random. Random sampling uses various techniques to ensure that the people who answer the survey are representative of the larger population being studied. The data obtained from the survey can then be corrected if necessary and used to make statistically valid estimates about the larger population. Surveys which can make statements about number of people in the U.S. who use the Internet or the WWW, for example, are using random sampling. Our survey uses non-random sampling which means we rely on users to see announcements of the survey in order to participate. Obviously, only those uses who see the announcements ever have the chance to participate. As a result, all segments of the Web population may not be represented in our sample. This reduces the ability of the gathered data to generalize to the entire Web population. At the heart of the problem is the fact that the Web does not yet have a broadcast mechanism nor a way of registering individual users (with digital signatures, for example) which makes it impossible to draw a random sample from a complete, or nearly complete, list of Web users. Over the course of the surveys we have used several methods to maximize the chances that our respondents do represent the larger web population and measure how well they do.

The first method we began using was promoting the survey through diverse media to attract respondents, including:

- links to the survey from high-exposure, general-interest Web sites, such as NCSA's "What's New", Yahoo, Lycos, CNN, etc.,
- announcements on WWW and Internet related Usenet newsgroups,
- coverage in national and local newspapers and trade magazines,
- an announcement on the *www-surveying* mailing list which we maintain for users who would like to be notified about GVU survey activities.

We felt that by providing many channels to bring respondents to the survey, we would attract a larger and more diverse set of users. To determine if the different channels were indeed attracting different sets of users, starting in the third survey, we have included a question asking how the respondent found out about the survey. This allows us to group respondents accordingly and look for differences between the different populations, specifically gender differences. For the third survey, we reported that there were no significant differences between the response profiles of women and men for the following categories: remembering to take the survey, other Web pages, the newspaper, other sources, and listserve announcements. There were differences found for: finding out via friends, magazines, Usenet news, and the www-surveying mailing list. Differences were even more pronounced in the fourth survey and we expect to find the same in the fifth. Given the low effectiveness of all but other Web pages and Usenet news announcements, which account for well over 50% of the respondents, most of these differences lead to nominal effects. To be conclusive, we would need to examine other basic demographics (e.g. age, location, income) across the different populations, as well. The differences in gender across the populations, however, are a positive indication that the different channels are reaching different sets of web users.

Another method we rely on is oversampling: collecting data from many more users that are required for a valid random sample. For the third and fourth surveys, we were able to collect data from approximately 1 out of every 1000 web users (based on current estimates of the number of people with web access). For random sample surveys, having a large sample size does not increase the degree of accuracy of the results. Instead, the accuracy depends on how well the sample was chosen and other factors [Fowler 1993]. Since we use non-random sampling and do not explicitly choose a sample, having a large sample size makes it less likely that we are systematically excluding large segments of the population. Oversampling is a fairly inexpensive way to add more credibility to a non-random web-based survey. The cost to actually collect data from extra users is minimal compared to other surveying methods; most of the expense is in the fixed costs of survey development and equipment and does not depend on the number of users surveyed.

When conducting a survey it is also valuable to know something about those who had the opportunity to respond, but did not. Ideally, we would like to know why they did not respond, but in most cases this is impossible. Instead most surveys simply measure the *rate* of non-response--the number of users who chose not to respond. For the third survey, we developed a similar measure of attrition rates. Attrition can best be thought of in terms of the paths taken by users through an information space. These paths are determined by the underlying structure of hyperlinks, that is, which pages are connected to which other pages. We know that some users will visit a page and not continue traversing the hyperlinks contained in that page. Others, however, will proceed to traverse the presented links, thus continuing down a path. Attrition for a particular survey can be understood as a measure of the percentage of users who began that survey, but who did not complete it. Attrition is calculated across a group of users. Attrition curves are defined as the plot of attrition ratios for all pages along a certain path. A complete discussion of the attrition analysis can be found in [Pitkow & Kehoe 95]. Excluding one questionnaire that had technical problems with submission, attrition rates for the third survey ranged from 4.54% to 12.58%.

Around the time that the fourth survey was completed, several other North American random-sample surveys released the results of their studies of Web and Internet users [Nielsen, FIND/SVP, O'Reilly]. These surveys used random-digit dialing (Nielsen, FIND/SVP, O'Reilly), an on-line questionnaire (FIND/SVP), and focus groups (FIND/SVP) to collect data on Internet and Web users. An obvious method of investigating the biases introduced by non-random sampling is to compare our results to theirs. The fourth survey's ratios for gender and other core demographic characteristics like income, marital status, etc., are almost exactly those reported by these other surveys. While our surveys do attract heavier Web users than do random phone-based surveys, it does not appear that frequency of Web use is a differentiating characteristic within the population. This result is both surprising and encouraging for web-based surveying.

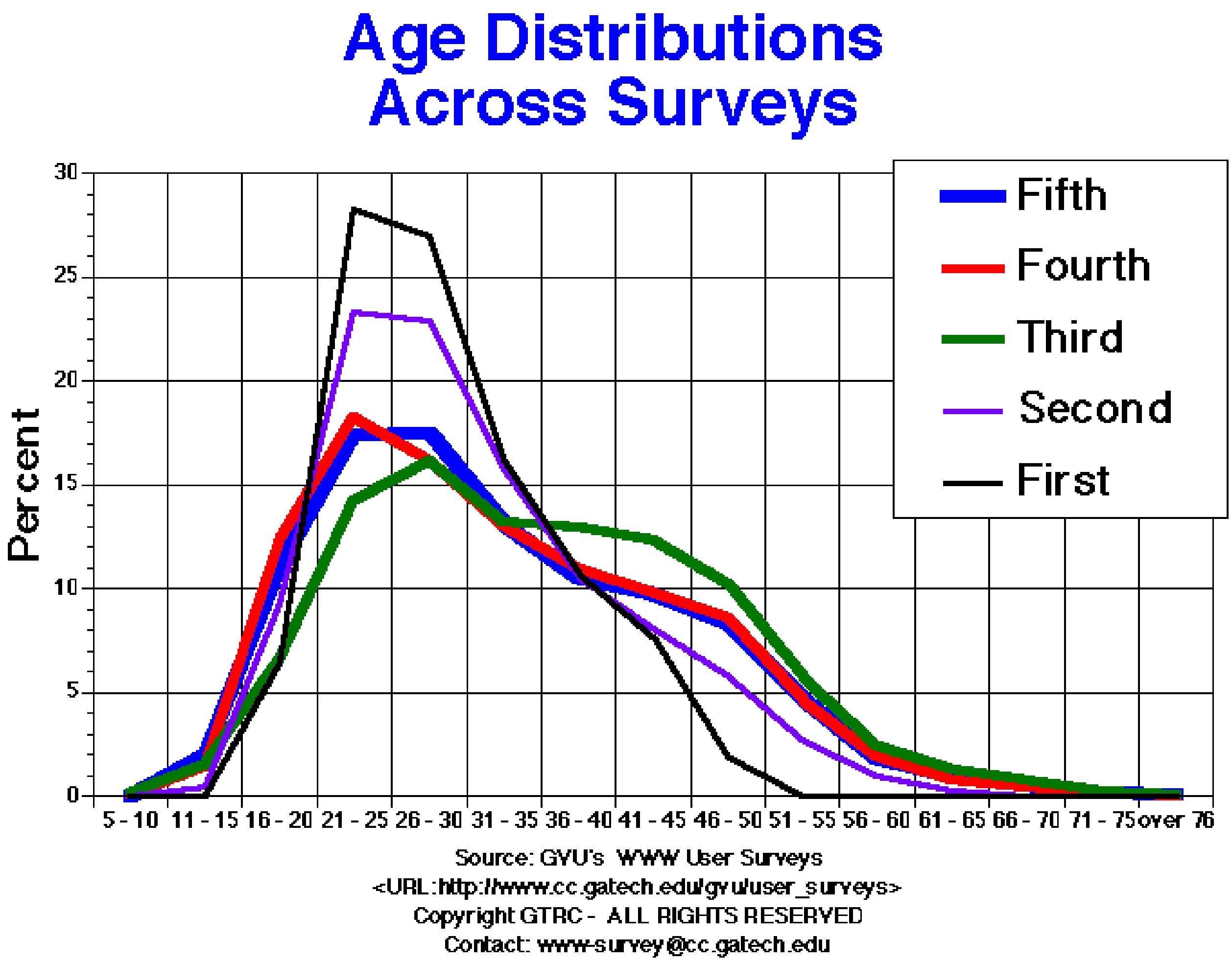
These methods when coupled with conservative interpretation of the data, lend a great deal of credibility to the results from the survey. One possible improvement that we are considering for future surveys is to select a random sample from the collected results. Data from other questions in the survey, such as the number of hours spent on the web, could be used to take into account the probability of selecting each person in the sample. Results obtained with this method could then be used to make statistically valid statements about the web population as a whole. Still, we remain unconvinced that the survey's sampling methodology is optimal and welcome suggestions and further comments on this subject.

Evolving Population

One of the most interesting aspects of studying the Web population is documenting the swift changes that it has gone through. While certain characteristics of the Web users sampled in the Surveys has remained the same or changed slightly, other characteristics have changed dramatically. More than ever, the users in the most recent surveys represent less and less the "technology developers/pioneers" of the earlier surveys (primarily young, computer-savvy users) and more of what we refer to as the "early adopters/seekers of technology." The adopters do not typically have access to the Web through work or school, but actively seek out local or major Internet access providers. As the Web continues to expand its horizon of users, we expect, and indeed find, that more and more users from diverse segments of the population participate in the Surveys. Please refer to the results from the individual surveys for more complete results.

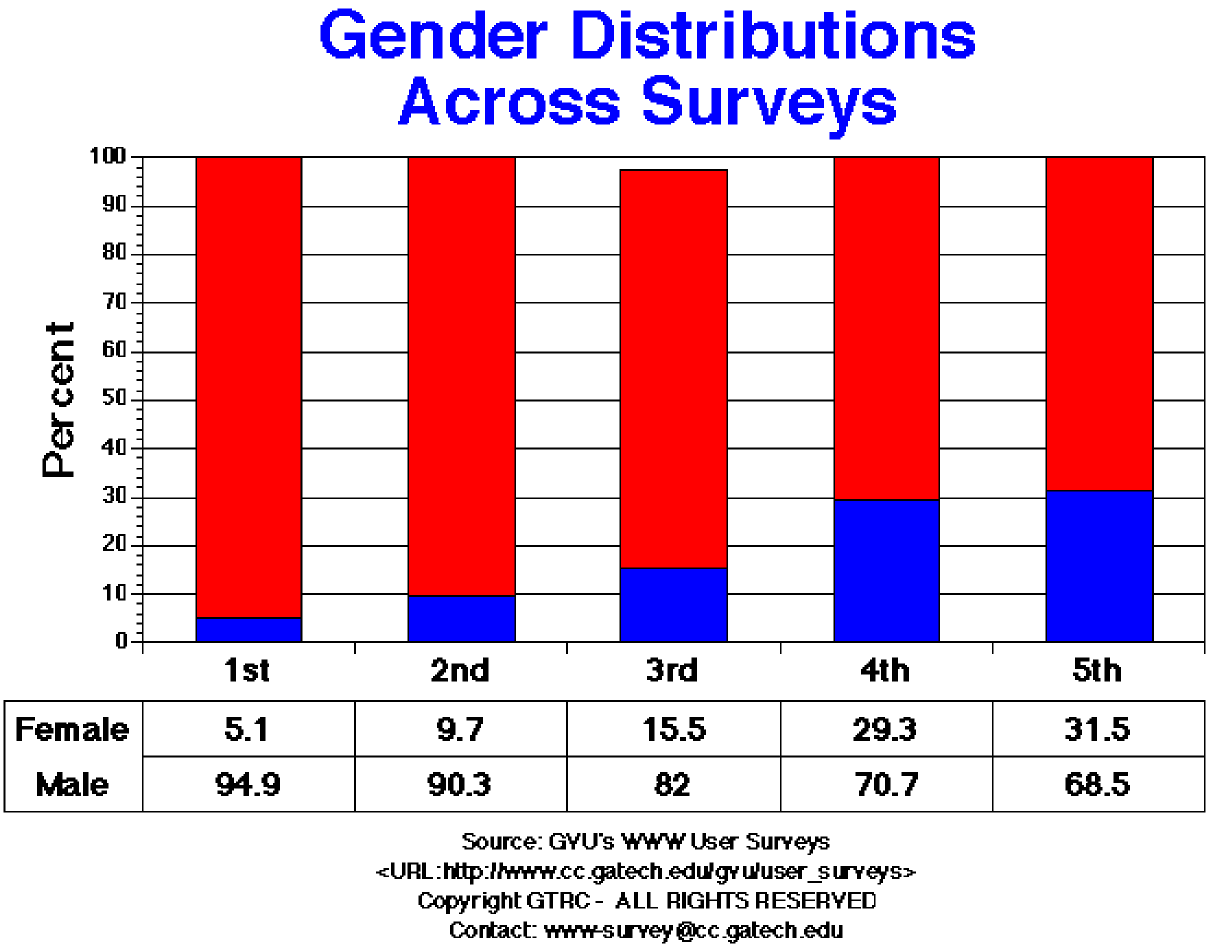
Age

The average age of respondents in the fifth survey is 32.95 years which is very close to the average from the fourth (32.7 years) and down two years from the third (35.0). Although the average age is relatively stable, we do notice dramatic changes in the age distribution. With each survey, the curve becomes flatter as more people in both ends of the age spectrum start using the Web.



Gender

The gender ratio continues to become more balanced with females representing 31.45% of the respondents and males representing 68.55% in the fifth survey. The percentage of females using the Web has more than doubled since the second survey (15.5% female, 80.3% male, 2.6% chose not to answer). Also, the US is integrating female users into the Web user population faster than Europe (US: 34.35% female, Europe: 15.2% female). The increase in female users is occuring largely in college students and K-12 educators.



Location

When classifying users by major geographic location, we find that the Web is slowly becoming less US dominated (US respondents: 80.6% third survey, 76.2% fourth, 73.5% fifth). Although Canada and Mexico showed a surge in the fourth survey (5.8% third, 10.2% fourth, 8.44% fifth), Europe has the second highest response rate in the fifth (9.8% third, 8.4% fourth, 10.82% fifth). All other areas of the world continue to show increases in the fifth survey including Oceania with 3.63% of respondents and Asia with 1.81%.

Education

Overall, the distribution of educational level has shifted slightly towards lower levels as indicated by less advanced degrees and more high school and some college level education. This trend towards more and more Web users without advanced degrees has continued since the second survey, where for example, over 13% of the users had doctoral degrees, compared to 7% for the third survey and 4% for the fourth and fifth. The education level of users is still high, in general, with over 80% of respondents in the fifth survey having at least some college education.

Primary Computing Platform

Unix, was the primary platform of most users in the second survey (44% second, 10.4% third, 8.76% fourth, 6.67% fifth) but some flavor of Windows has held this position since the third (29% second, 51.98% third, 61.5% fourth, 63.63% fifth). The Macintosh platform has accounted for between 20% and 30% of the users in each survey.

Years on the Internet

There seems to be a fairly steady stream of new users to the Internet as indicated by the percentage of users who have been on less than twelve months: 50.2% for the third survey, 60.3% for the fourth, and 43.14% for the fifth. The rise in the number of new users in the fourth survey can probably be attributed to users who have gained access through local online providers.

Nature of Internet Provider

The nature of respondents Internet provider has shown substantial change throughout the surveys. (There was a link from Prodigy to the Third GVU survey, so results from that survey are probably biased for this question and are excluded from this analysis.) The percentage of users gaining access through educational institutions has dropped from 51.0% in the second survey, to 31.59% in the fourth, to 26.8% in the fifth. The most popular method of gaining access in the fourth and fifth surveys is through local Internet providers (41.64% and 48.53%, respectively) while access from major providers accounts for only 8.1% in the fourth and 9.24% in the fifth.

Willingness of Users to Pay for Access

One of the most stable characteristics of the earlier surveys has been that one of five users stated outright that they would not pay for access to WWW sites. This number has increased from 22.6% in the third survey, to 31.8% in the fourth, to an amazing 64.95% in the fifth. This is indeed alarming for those who wish to apply a subscription business model to the Web. This may also very well reflect the perceived value of the material and resources currently available on the Web by its users. It may also be related to the fact that 57.64% of the users in the fifth survey are paying for their own Internet access. For those who were willing to pay, the largest percentage (12.06%) favored a subscription model.

Conclusion

Measuring and describing the Web population has turned out to be an interesting and challenging task. A primary goal of ours has always been to provide quality data, with the limitations clearly defined, at make it available to support a variety of research agendas within the Web community. We feel that through our technology and methodology, we have been able to reach this goal. We hope that as more researchers enter this field, new ideas and collaborations will continue to raise the quality of the data being collected.

Footnotes

Rule based adaptive questioning and enforced completion were implemented in January 1994 by John

Mallery for an email based survey of White House document readers. (See Roger Hurwitz and John Mallery, Of Public Cyberspace: A Survey of Users and Distributors of Electronic White House Documents, http://www.ai.mit.edu/projects/iiip/doc/surveys/report.html .) It was at Mallery's suggestion at the Second World Wide Web Conference that these features were incorporated into the GVU survey.

References

Alao, F. (1994). Pilot Study of Network Surveying Techniques. (manuscript not published).

Catledge, L. and Pitkow, J. (1995). Characterizing Browsing Strategies in the World-Wide Web *Journal of Computer Networks and ISDN systems*, Vol. 27, no. 6.

FIND/SVP. (1995). The American Internet Users Survey. http://etrg.findsvp.com/index.html

Fowler, F. (1993). Survey Research Methods, Second Edition. SAGE Publications, Newbury Park.

Nielsen Media Research/CommerceNet. (1995) The CommerceNet Nielsen Internet Demographics Survey. http://www.nielsenmedia.com/

O'Reilly Research. (1995). Defining the Internet Opportunity. http://www.ora.com/survey/

Pitkow, J. and Recker, M. (1994). Results from the First World-Wide Web Survey. *Journal of Computer Networks and ISDN systems*, Vol. 27, no. 2.

Pitkow, J. and Recker, M. (1995). Using the Web as a Survey Tool: Results from the Second World-Wide Web User Survey. *Journal of Computer Networks and ISDN systems*, Vol. 27, no. 6.

Pitkow, J. and Kehoe, C. (1995) Results from the Third WWW User Survey, *The World Wide Web Journal*, Vol. 1, no. 1

Acknowledgments

Georgia Tech's Graphics, Visualization, & Usability (GVU) Center operates the surveys as a public service as part of its commitment towards the Web and Internet communities.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Thanks to all members of the GVU, its director Dr. Jim Foley, and staff for their support and help. Special thanks extend to Kipp Jones, Dan Forsyth, Dave Leonard, and Randy Carpenter and the entire Computer Network Services staff for their technical support and Sun Microsystems for their generous donation of equipment.

Author Information

JAMES PITKOW received his B.A. in Computer Science Applications in Psychology from the University of Colorado Boulder in 1993. He is a Graphics, Visualization, & Usability (GVU) Center graduate student in the College of Computing at Georgia Institute of Technology. His research interests

include event analysis, user modeling, adaptive interfaces, and usability.

COLLEEN KEHOE received her B.S. in Computer Science from Stevens Institute of Technology in Hoboken, NJ in 1994. She is currently a Ph.D. student in the Graphics, Visualization, and Usability Center of the College of Computing at the Georgia Institute of Technology. Her current interests include educational technology, visualization, cognitive science and Web-related technologies.