Web Usability: The Search for a Yardstick (Panel)

Organizer: Jean Scholtz, National Institute of Standards and Technology

Panelists: Amy Schwartz, Ameritech
Andrew Sears, DePaul University
Coco Becker, Intel
Laura Downey, NIST

Abstract

The design and evaluation of World Wide Web sites present interesting challenges for human factors and usability professionals. One of the most challenging aspects is the speed at which development of Web sites occurs. As more and more information appears on the World Wide Web, the complexity of access and use is greatly increased. It is essential that we identify usability criteria for the Web along with tools and methodologies for design and evaluation that will enable web developers to produce and maintain good web sites.

We are interested in how Web site designers and maintainers design and evaluate their Web sites. Our panelists have all been involved in the design and evaluation of Web sites. We have asked them to share with us their techniques for assessing the usability of their sites by replying to the following three questions:

- What are the top two criteria you use in designing and developing Web sites?
- How have you tested these criteria during design and development?
- What other tools or techniques would have helped you in the process?

Our goal is that the body of information collected during this panel not stop with the concluding remarks of this panel, but rather be the start of a collection of problems and possible solutions. We encourage conference attendees to publicize this site to their colleagues and friends so that we may share best practices and make the Web a better place to browse.

Introduction

Development and use on the World Wide Web are increasing at exponential rates. The Zeus Web site reports that over 20 million people worldwide now use the World Wide Web. Brown Computer Solutions shows that the number of Internet hosts has quadrupled in the last year. The advantage of having all this information at our fingertips may soon disappear if we find ourselves buried in information. Instead of being buried in paper, we'll be buried in bits! It is imperative that researchers and developers quickly turn more

attention to ensuring that web sites are usable.

Currently, we use several types of tools and techniques to produce usable graphical user interfaces (GUIs) including: interactive software design, GUI guidelines, and numerous usability testing techniques. Are these tools and techniques appropriate to use in the design and development of Web based applications? Are the usability criteria we're measuring and evaluating for current software designs what is needed for Web applications? Given the sudden rash of articles, tutorials, workshops, and discussions we've seen of late, we think not. Why is that? What is it we need instead? And how are we going to get what we need QUICKLY!

Participants in the workshop on Web usability at CHI'97 expressed the need for rapid, remote, and automated tools and techniques to achieve this goal. While the speed of development on the Web is a challenge and constraint we must deal with, the Web holds many possibilities for tools and techniques to use in design and in evaluation, not typically available in traditional software environments. The possibility of being able to automate data collection and to collect data from many remote users is now feasible. What type of information can be collected and how is it useful in determining if the site is meeting usability criteria?

Development Differences

What are the possible differences in Web development and use that make it somehow different from traditional software development? Speed of development is potentially the most visible difference. Currently, the types of sites and applets that appear on the Web are simpler than many traditional software projects. In many instances, Web development is done incrementally. Some developers even consider the initial releases as 'beta tests" and provide a mechanism for collecting data to use for successive versions. Is this type of feedback useful and valid? What types of information should be collected in this fashion? Is it necessary to do traditional usability testing prior to putting out software for the public to use? And, if so, exactly what types of testing can be done in the Web development time frame?

Rapid changes in development tools also affect Web software. Every day new technologies allow developers to incorporate more or better capabilities in their application. Multimedia techniques continue to improve, allowing access to audio, video and animation. As more of these capabilities appear, developers attempt to take advantage of more and more. Guidelines for GUI development are still being developed in these areas. Those that address these issues on the Web are still evolving. The capabilities appear and get pressed into use before assessments have been made about user effects.

Web usability criteria

Traditional graphical user interfaces are evaluated with respect to the following criteria: (Shneiderman, 1989)

- Time to learn
- Speed of performance

- Rate of user errors
- Retention over time
- Subjective satisfaction

Are the same criteria applicable to Web based applications? In GUI development, usability criteria are given different weights depending on the type of application and type of user base for that application. Can we predict the type of user base and the way in which applications will be used on the Web?

Traditional software applications are purchased by users (or nowadays, downloaded). Packages and advertisements spell out what the software does and set users expectations in that way. How do we set user expectations about Web applications?

Moreover, when purchasing traditional software products the brand name is prominently displayed on the packaging. Is branding still important on the Web? And if so, how do companies achieve branding?

Panel Position Statements

We have assembled a panel of experienced Web designers, evaluators and maintainers who have worked with a variety of Web sites. Our goal is to look at various criteria that are currently being used, the methods that HCI professionals are using to evaluate the sites with respect to those criteria and their suggestions for other tools, techniques that would facilitate design and development. We've asked them to respond to three questions:

- What are the most important criteria you use in designing and evaluating your Web sites ?
- What are the tools and techniques you use during design and evaluation and how do these relate to the criteria?
- What are the highest priority items on your wish list for help in design and evaluation of Web sites?

Amy Schwartz, Ameritech

The top two criteria we use are:

- Is the content organized in a logical and use-sensitive manner with clear navigation?
- Is the interface easy enough to use without documentation or training?

We use usability testing, search term analysis, card sorting techniques, focus groups, log analysis, and feedback form analysis to test and refine our web sites.

We rely on usability testing to help assess the overall ease of use once we have a

prototype to put in front of users. Usability testing really helps us find the big problems with the interface, especially in terms of navigational difficulties, confusing terminology and user expectations about how the interface will work.

Search Term analysis gives us insight into the goals that users have for visiting the site and the content they would like to see. We keep a log of all the search terms entered on our external corporate site and have classified according to the goals we infer they represent and according to the content for which people searched. We also get the top ten search terms used in our site each week. We compare these terms from week to week to determine if these are the most popular search terms or terms that are difficult to find.

Card sorts help us determine how users expect our Web site content to be organized. Pre-launch focus groups are also useful in determining what users expect to find. Post-launch focus groups are helpful in discovering navigation problems, and learning what content and capabilities users wanted but weren't finding. The post-launch groups were also helpful in identifying the tasks for which the users actually used the site.

We use analysis of Referer logs to tell us where people are coming from and where they end up. From this we can get a good idea of users' goals for coming to the site. Our transfer log provides us with hourly, daily, and monthly page request information for individual sections of our site. It is helpful in determining when and how often certain sections of the site are being accessed. With this sectional information we do a separate analysis that compares the sections to each other in whatever category (hourly, daily, monthly, or even total page views) we desire. This can prove to be important in determining the effectiveness of an ad, for example, in a site.

Our external corporate site has feedback forms on virtually every page. The questions asked on these forms are logged and classified according to the goals we infer they represent and according to the subject about which people ask.

Other tools or techniques that would have been helpful include more people, a better statistical package that graphs, in HTML, pie charts of custom referrer log analyses, feedback forms that require users to classify the type of question they are asking, and more information on client configuration.

Andrew Sears, DePaul University

After designing and evaluating Web sites, and conducting research into a variety of issues related to Web site development, it is clear that developing for the Web is as complex, or even more complex, than designing traditional computer applications. Traditional issues persist, but new issues also become apparent. Two specific issues that require particular attention when designing for the Web are: response times and navigation.

While response times have been studied extensively, distributed applications raise new concerns. Delays are longer, more variable, and are often beyond the control of individual users or organizations.

One common complaint is that documents take too long to download. We need to re-open the investigation into response times. We need to understand the characteristics of delays

that are introduced by networks and how these delays impact users. Using this knowledge, we can design more effective documents and more effectively manage user expectations.

Navigation is an obvious problem. The quantity of information available on the Web continues to increase dramatically. While the answer to your question may exist, it is often hard to find. This is true even within a single Web site. How do users navigate through the vast quantity of information available to find the answers to their questions? Traversing links can work, if we know the questions ahead of time, but that is not always the case. Searching works, if users understand how to use the search facilities and know the right keywords. How can we provide an interface that accommodates users that use both of these strategies?

As navigation and response time issues impact all Web users, I propose that two general criteria for Web sites are:

- Design effective documents and effectively manage user expectations so that delays in response time or variable response times do not adversely affect users.

- Design Web sites that support both searching behavior and link traversal by users.

Unfortunately, there seems to be a lack of quality tools to assist in the process of evaluating these problems. At the same time, current guidelines are often inadequate. Network delays are complex and vary from site to site. Delays vary depending on the location of the end user, the time of the day, the day of the week, and a variety of other issues that are also beyond the control of the user. Due to the lack of existing software to deal with response time issues, we have developed a suite of tools that I use. This suite of tools can help in a variety of ways, but I will focus on two.

- Using an instrumented server, documents can be delayed based on the network conditions being simulated and the documents being retrieved. This provides evaluators, for the first time, with a means of viewing their documents while experiencing the delays users will experience.

- A second tool automatically identifies portions of the site, documents, or subdocuments that may lead to response time problems. This is done by loading the documents repeatedly and identifying any documents that are consistently slow.

At present, I believe there are even fewer tools that help designers deal with navigation problems. Navigation seems to rely on so many factors that this reverts to basic interface design and understanding what your users want to accomplish. Navigation isn't a question of whether a link exists, it's a question of whether the user will find, recognize, and select that link.

My wish list includes a set of tools that can automatically test a variety of issues. Standard HTML and link checkers are a start, but are inadequate for really evaluating usability. Tools that deal with response time issues are currently being developed. Being able to provide developers with a realistic view of their documents (including the

network delays) seems like an important step in the process of developing tools for Web developers. These tools must be useful in a development process that moves quickly and never really ends. Web sites also depend on sites developed by other organizations, which leads to even more problems. How do we know if someone else has inadvertently broken our site (e.g., moving a document that we reference)?

Coco Becker, Intel

For us, the important criteria for our Web site include:

- Intuitive navigation for a wide audience of Intel products and technologies (developers, end users, investors, job seekers, channel, distributors, etc.). With everyone from software developers wishing to enhance their applications to include MMX (r) technology to small investors wishing to purchase Intel stock, our web site must meet all of these diverse user requirements.

- Enhance the Intel brand. We are a company that represents innovation, technology as well as high quality. Our Web site needs to reflect and reinforce that image in every element. That includes everything from consistent look and feel of each Web page all the way through outstanding server performance to eliminate downtime.

- Compatibility with a wide range of desktop configurations. Our audience base consists of Win 3.1, through Win NT users, UNIX users and 14.4 dial-up access through to dedicated T1 access. We have a mix of browsers and browser versions. One of our design requirements is that the user experience is consistent as much as possible across all platforms and plug-ins.

How do we measure ourselves?

- We do extensive usability testing when launching new sites or new redesigns.

- We track many statistics, which we make available to authors on a daily basis via an Intranet Web site. These statistics include:

    - Browser version customers use when coming to our site.

    - Adherence to corporate look and feel guidelines (proper header and footer usage, meta-tags, etc.)

    - Customer feedback to our Web site - both volume and category of feedback.

    - Customer usage, frequently entered key words into search engine.

- Monthly review with a management review committee to track progress.

What do we wish we had?

- Scalable tools to handle large volume of log files. Most tools we have seen can

only store 1-2 days worth of our log files.

- Better tools around managing "content templates" - manipulating values in meta-tags, etc. For example, when a person moves to another job, going through all of their files and updating the author tag to the new owner.

- Compatibility testing support - especially around browser versions and different dial-up speeds.

Laura Downey, NIST

As a research computer scientist/usability engineer at NIST, I have been called upon to evaluate Web sites at various stages and also to design Web sites. The majority of Web sites at NIST are concerned with providing information. In other words, the focus is on content and not task. This I believe is true for much of the Web, but may change as more transaction processing is performed on the Web and other task-based activities are introduced. From this content perspective, I usually develop a set of usability goals with the site owner. These goals are based on information location/retrieval of the content. The other criteria I establish involves consistency in terms of site identity which is done by developing a page template involving contact information, layout, logos, etc.

In terms of evaluation, most often I perform expert reviews based on usability goals and/or standard heuristics (such as Nielsen). However, when possible, we take at least two basic measurements during usability testing. One, a binary (yes/no) measurement against the usability goals, and two, if the goal is met, time to locate the information. This performance time is then compared to an "expert" performance time. I have often wished that there were more tools available to do things such as consistency checking and user logging. At NIST, we are investigating developing tools to analyze server log data, instrumenting client actions and analyzing HTML structure. The recent popular press indicates that some tools are being developed. However, from most corporate perspectives, they do not address "their" issues so they are developing their own set of tools as best they can. What we need are ways to rapidly design and evaluate sites since the Web moves so quickly. Automated tools can help us do this. It would also be useful to have a cookbook of usability measurements for the Web.

Panel Outcome

The outcome of the panel will be a prioritized list of criteria used in Web design and evaluation along with descriptions of the methods currently being used to assess those criteria. In addition, we will compile a list of desired tools/ techniques that would be beneficial to the web development community. This information will reside on a Web site maintained by NIST. Members of the community will be invited to use and to contribute to this information so that it becomes a repository of best practices for Web criteria, Web design tools, Web usability methods, and Web usability tools.

Last Update: June 4, 1997