Alpert, S.R., Karat, J., Karat, C-M., Brodie, C., & Vergo, J.G. (2003). User Attitudes Regarding a User-Adaptive eCommerce Web Site. *User Modeling and User-Adapted Interaction*, *13*(4), 373-396.

User Attitudes Regarding a User-Adaptive eCommerce Web Site

Sherman R. Alpert, John Karat, Clare-Marie Karat, Carolyn Brodie, John G. Vergo

IBM T.J. Watson Research Center
PO Box 704
Yorktown Heights, NY 10598 USA
{salpert, jkarat, ckarat, brodiec, jvergo}@ us.ibm.com

ABSTRACT

Despite an abundance of recommendations by researchers and more recently by commercial enterprises for adaptive interaction techniques and technologies, there exists little experimental validation of the value of such approaches to users. We have conducted user studies focussed on the perceived value of a variety of personalization features for an eCommerce Web site for computing machinery sales and support. Our study results have implications for the design of user-adaptive applications. Interesting findings include unenthusiastic user attitudes toward system attempts to infer user needs, goals, or interests and to thereby provide user-specific adaptive content. Users also expressed equivocal opinions of collaborative filtering for the specific eCommerce scenarios we studied; thus personalization features popular in one eCommerce environment may not be effective or useful for other eCommerce domains. Users expressed their strong desire to have full and explicit control of data and interaction. Lastly, users want readily to be able to make sense of site behavior, that is, to understand a site's rationale for displaying particular content.

KEYWORDS

adaptive interaction, personalization, eCommerce, user profile, user studies, human–computer interaction, collaborative filtering

1 INTRODUCTION

Personalization is the use of information about a particular user to provide tailored (personalized) user experiences for *that* user. In the context of the World Wide Web, personalization involves automatic customization of Web pages to accommodate individual users' needs, interests, knowledge, goals, or tasks. A personalized Web site can be described further as a system that "adapts the content structure, and/or presentation of the networked hypermedia objects to each individual user's characteristics, usage behavior, and/or usage environment" (Kobsa, Koenemann, and Pohl, 2001, p. 3).

Personalization may encompass a wide assortment of technologies in addition to user-adaptive facilities. In the context of eCommerce, personalization might also include customer relationship management (CRM) mechanisms (for example, integrating purchase histories from all customer touchpoints, including Web site, brick-and-mortar stores, and call centers) or providing for a "personal shopping experience" (for example, providing immediate contact with a live technical or sales representative who is aware of the customer's current context on the site). In this paper we focus only on the adaptive interaction aspects of personalization.

Adapting content and interaction based on knowledge about users is certainly not a recent innovation. Since the 1970s, knowledge-intensive user models have been used by researchers and developers of intelligent tutoring systems to guide individualized instructional interactions with learners (see, e.g., Sleeman & Brown, 1982). Beginning more than two decades ago, researchers involved in advice-giving systems and natural language generation have described applications that would tailor text to individual users or user stereotypes (e.g., Rich, 1979; Paris, 1987; 1993). More recently, user-adaptive technology has been applied to a variety of domains including education, information retrieval, and virtual museums, primarily on the Web (Brusilovsky, 2001).

Now researchers and developers are applying and adapting the work of earlier adaptive system research and solutions in other domains to the particular problems and requirements of eCommerce Web sites. The basic tasks for eCommerce users are researching and learning about products and services that most closely match their needs, and actually purchasing such products and services. These fundamental activities may involve or spawn

other informational and purchasing tasks, such as obtaining product support, purchasing add-ons for previously purchased products, and so on. Specific suggestions from the research community regarding tailored interaction for eCommerce have included natural language generation for personalized product catalogs (Milosavljevic & Oberlander, 1998) and using previously monitored user behavior to infer that user's needs and preferences in order to present products and product descriptions adapted to those requirements (Joerding, 1999; Ardissano & Goy, 1999). There are numerous commercial software products targeted at enterprises that wish to implement personalization in some form on their eCommerce site and that incorporate user-adaptive features toward that end (e.g., IBM WebSphere® Personalization, BEA™ WebLogic Personalization Server™, BroadVision One-To-One® Enterprise™, many others). There exist numerous consulting firms whose raison d'etre is to help companies add personalization to their sites. And scores of extant Web sites incorporate personalization features.

There are serious questions that must be asked in the face of this background. Do adaptive systems actually improve the user's experience when using such systems (see also Chin, 2001)? What particular adaptive features improve and which features detract from user experiences on eCommerce Web sites? Does personalization actually add value to users of an eCommerce site and thereby encourage them to return to the site? The other side of the equation is the enterprise hosting the personalized site—there may be numerous benefits to this enterprise as well as its customer. The simplest is that a customer who is happy with her user experience on an eCommerce site may wish to return again and purchase more. But here we will focus on the customer/user perspective.

We have been involved in the user-centered design of personalization for eCommerce. The user-centered design approach bases design decisions on the experimentally validated value *to users* of application features and tools (Vredenburg, Isensee, & Righi, 2001). We have conducted user studies whose initial and primary purpose was to determine which specific personalization features would be judged the most usable, valuable, and attractive to users of an eCommerce Web site. Thus, initially, the user studies described herein were not performed to investigate user attitudes toward user-adaptive sites per se! We gathered a large amount of quantitative data toward our initial goal. Our experimental design allowed us to obtain a wealth of qualitative data as well, in the forms of

written and spoken opinions of design concepts and interaction scenarios. This data included the results of written questionnaires, think aloud protocols, and free-form group and one-on-one discussions, as well as our own observations of users in a laboratory setting. The data gathered allowed us to successfully derive a set of core personalization features (see Karat et al., 2002). From this well of data also emerged clear attitudes of users toward adaptive techniques that were intrinsic to the implementation and design of the personalization features we tested. This paper focuses on those attitudes, opinions, and reactions.

2 METHODOLOGY AND USER STUDIES

We executed an iterative series of three studies, carried out in multiple laboratory settings. The studies were a mixture of group sessions in which participants were walked through the use of a prototype by an experimenter, and individual user evaluation sessions in which subjects interacted with a working prototype. The facilities of the usability laboratory enabled observation of the user session in the "Studio" from the "Control Room" through use of a one-way mirror and video monitor, as well a range of data collection activities.

2.1 THE PERSONALIZATION FEATURE SPACE

Personalization and user-adapted interaction cannot proceed without obtaining information about the user. Thus our initial step was to gather information from as many sources as possible about techniques that might be used to provide user information to the system. We then added to this feature list personalization features that *use* information about the user to adapt interaction and the information displayed for users to view. Our information gathering included literature searches along with brainstorming sessions with a number of groups within our Research Division working on technologies to understand user context (even to the extent of inferring user mood). At this point we did not want to prejudge techniques for lack of feasibility.

We initially catalogued 75 techniques (space prohibits us from providing the complete list here, however). We clustered the techniques into categories based on similarity as shown in Table 1. We then began to prune and focus this list based on our own heuristic analyses and measures of current implementation feasibility. For example, we decided to eliminate as a feature to be tested the use of the inferred mood of the user as a basis for adaptive content presentation. We were careful to ensure that we included features from each category. The goal was ultimately to refine the list based on measures of effectiveness,

usefulness, usability, and user attitudes derived from successive user studies. User-adaptive features and techniques form a subset of the full complement of personalization features initially considered and those ultimately tested in the user studies.

PersonalBook – a distinguished place on a site where all personal data can be accessed and modified
Universal profile – a single place where user information is persistently stored; used by all functions on the entire site
Subscription-based services
Personalized service and support
Recommendations based on profile data
Adaptive presentation tailored to user characteristics
Personal Preferences in page layout or format (customization)
Adaptive navigation
Live (chat-like or phone-based) help or sales support (Personal Shopper)
Personalized feedback, such as indication that user is recognized, personalized messages
Accessible transaction history
Loyalty programs, incentives
Future purchase considerations
Your personal store, built by an expert

Table 1. Personalization Feature Clusters

2.2 MATERIALS

2.2.1 Prototype Adaptive Web Site

We then constructed prototypes of an eCommerce Web site. There were two prototype systems used in the user studies. Both were implemented in Microsoft PowerPoint and presented on an IBM ThinkPad computer. In the first two studies, the low-fidelity prototype consisted of a series of screen shots; when the lead experimenter virtually "clicked" on a widget, the next screen shot—the site's "response" to the virtual click—was displayed. After the first two studies, we created an interactive prototype for the Study 3. Although PowerPoint is well known as presentation-only software, the latter PowerPoint program was a fully interactive application that accepted user input and acted based on such input.

The prototypes were designed to demonstrate specific personalization and adaptive features. The system exemplified in the prototypes is a Web site that maintains a profile of

the user's personal information and tailors the site's content to that user based on profile information and navigational context.

The central personalization tool used in our scenarios is the *PersonalBook* (see Figure 1). It is closely tied to the user's profile, the persistently stored personal information about the user. The PersonalBook can be opened or closed at the user's will (as well as programatically as appropriate). Once displayed, the PersonalBook provides several informational and task-oriented tools. For example, the user may maintain a list of all products that she or her organization owns (see Figure 2). A click on any product listed in "Products That I Own" brings the user to a page showing those accessories and upgrades that are compatible with the selected product as well as other information specific to that product. This provides a fast navigation path to finding appropriate add-ons for products already owned.

Another tool in the PersonalBook, "Help me find what I need," attempts to help users locate products and services on the site. This tool is context-sensitive, using information about the currently visible page to guide its content. For example, if the user is viewing a page showing the full set of server families when he opens the "Help me find what I need" tool, its content includes a context-specific question, "What type of server are you looking for?", and questions also specific to the context, in this case "servers" (see Figure 3). The user may click on selections in any or all of the response categories, and by doing so specifies the criteria for a new product search. The site responds with a page whose content has been filtered based on these criteria—for example, the resultant page in this example may contain a single server family that matches the user's selections.

Other portions of the PersonalBook may be consulted for other eCommerce-related tasks or information, such as obtaining a user's full purchase history on the site, viewing named shopping carts that the user has saved for later retrieval, and so on.

The prototype, by definition, did not fully implement all of these facilities in all possible contexts. Rather, pieces of each are implemented and enabled in specific usage contexts so participants could complete scenarios in one of several implemented ways. The prototype provided enough functionality, and in the appropriate situations, that study participants could get a true feel for these features and provide realistic assessments thereof.

2.2.2 Written Questionnaires

The research team developed questionnaires for subjects to fill-out after viewing (Studies 1 and 2) or interacting with (Study 3) the prototype for each scenario. The questionnaires allowed us to capture both quantitative and qualitative data. Post-scenario questionnaires asked participants if they would use the prototype Web site to perform transactions similar to those shown in the scenarios, and asked them to explain their response. These questionnaires also asked subjects to rate the personalization features demonstrated in the scenario. Features were stated as assertions, such as the following:

- A personal "myXX" [my <name of company>] site is created for you when you provide information about yourself.
- You control all the data kept in your profile and can review and edit it at any time.
- You can create and use a custom catalog on the site that is tailored to the products you are interested in and those that you own.
- The site will conduct constrained searches for accessories and upgrades compatible with the products listed in your profile.
- Implicit data and information are used to create your profile.
- You have the choice of having search constrained based on current activities and profile data.
- The pages displayed are adapted based on your navigation path (e.g., the site remembers the search results and presents accessories appropriately when you move from notebook to desktop system).
- The pages displayed are adapted based on automatically identified, non-constant information, such as your connection speed (e.g., pictures are not displayed at low connection speed, but a link is provided to access them).
- You see the most popular things that other people like you have bought related to the item you are now purchasing.
- You can use the "Help me find what I need" function to help you filter through product choices and make purchase decisions.

In the first study, participants were asked to rank order the features shown in each scenario based on their value to the participants. We changed this metric for Studies 2 and 3: participants were asked to rate each feature in terms of its value to the participant using a 7-point scale ranging from "Highly Valuable" (1) to "Not at all Valuable" (7).

A post-session questionnaire was also administered after the last scenario. This questionnaire dealt with marketing and business case issues, including whether subjects thought they would be more likely to visit the eCommerce site and more likely to purchase from the site if the features of highest value were implemented (see Karat et al., 2002).

2.2.3 User Task scenarios

The research team then created a set of task-based scenarios that would provide a real usage context for participants as they viewed and/or interacted with the prototypes in the studies. The scenarios covered purchasing I/T equipment, maintaining and upgrading the equipment, and getting support for products. The user scenarios featured Pat User (whose gender was randomly assigned in each session) who needed to complete the set of tasks that arose across a period of 18 months in Pat's organization. Thus we examined both initial visits to a site with personalization features as well as repeated-use scenarios on the same site.

In the first two studies, participants were walked through three usage scenarios. In the third study, participants adopted the role of the user named Pat User and actually "drove" the interactive prototype. For this last study, we refined the larger scenarios into six shorter ones, each of which could be accomplished in ten minutes or less. Each scenario involved one or two tasks to be accomplished, and exercised—introduced in a concrete fashion to users—three or four personalization and adaptive features.

Scenarios were presented to participants as narratives, such as:

> Within the next year you plan to replace five of your department members' desktop displays with 22-inch LCD high-resolution flat panel displays. One evening you log on to the [vendor] site from home over a slow dial-up connection. On the site you look for displays that are compatible with the department members' NetVista computers. Find the 22-inch flat panel displays and save them in a shopping cart marked as future plans...

Two additional scenarios from Study 3 are presented below in order to portray the adaptive features explored by each. These scenarios are included to illustrate the interaction path participants followed in performing their assigned tasks.

2.2.3.1 "Compatible Memory" Scenario

"You and your department have made various server, laptop, and desktop purchases. You now think you may have to purchase additional memory to enhance to capabilities of the laptops used by your department members. Starting from your PersonalBook, find 128MB add-on memory chipsets compatible with those laptops. Then also find memory compatible with the desktop machines your department owns…"

To accomplish these goals, a participant would begin by opening the PersonalBook and the "Products That I Own" section therein (see Figure 2). She could then select the ThinkPad A30, thereby navigating to the compatible-accessories page for the specific computer model owned by her department—this page provides access to only those accessories and upgrades that are compatible with the ThinkPad A30. The user's task then requires either clicking on "memory" (in the table at the top of the page) or using the search field to find memory options, and in response the system displays a list of memory compatible with the ThinkPad A30 (as in Figure 1). This is an example of the adaptive technique of information filtering (or content filtering) wherein the content shown to the user is a subset of the full information available to the system, and that subset is based on information about the user. In this case, the full complement of all available memory chipsets is filtered so as to show only those compatible with the user's computer.

To continue the scenario, the user chooses the appropriate quantity of 128MB chip sets and places them in her shopping cart. Then the user returns to "Products That I Own" and selects the desktop computer listed therein. This time around, when the compatible-accessories page is shown, it now automatically displays memory options, with no further search or selection required. Here the site is making an assumption based on previous interaction: because the user had previously looked at memory compatible with the laptop listed in "Products That I Own," he might now be looking for memory compatible with the desktop system he subsequently selected. The content of the current page is thereby being adapted based on a recent navigation path: the site is using implicit information about the user to infer his current intentions.

2.2.3.2 "Purchasing a Server" Scenario

In this scenario, the user tries to find a server that meets specific performance and financial requirements, as stated in the scenario description in terms of the number of clients it must support, the type of application functionality the server will provide, and price range. Goal attainment is effected by use of the "Help me find what I want" wizard in the PersonalBook. "Help me find what I want" attempts to help the user navigate quickly to the product or service she in interested in researching. As described above, the specific question posed to the user and the widgets made available for the user's response are context-dependent.

To complete this scenario, a participant might navigate to the main server page, which displays four server families (as shown in Figure 3). To then narrow his search, the user can use the "Help me find what I want" wizard to filter the list of servers to those that meet the stated requirements; the resultant page displays only servers that match his criteria. Here the system supports the user in attaining his goal using information he explicitly provided.

2.3 PARTICIPANTS

Participants for the preliminary Study 1 were employees of the same company as the research team, recruited via broadcast messages sent by the team. In Studies 2 and 3, subjects were employees of external companies, recruited by an external vendor through use of user profiles (or "screeners") that specified characteristics of our target users. In Study 1, the research team similarly screened subjects. In the Results and Discussion section below, we focus primarily on the results from Studies 2 and 3 which were executed using external participants.

All study subjects were drawn from the population of people who are comfortable with the World Wide Web (3+ hours per week usage) and who are at least moderately "tech-savvy" in their purchasing behavior. That is, they regularly use the Web for both researching products and purchasing them, they make purchases on eCommerce sites themselves, but may enlist assistance from other technical experts in the selection of technology to be purchased. Participants had input to the decision process for the purchase of a server or workstation in the last year. And about half of them had also been involved in the purchase

decisions for desktop and notebook systems. Participants in Studies 2 and 3 received $150 for taking part in a two-hour user session; participants in Study 1 received a voucher for a free lunch in the company cafeteria. New groups of target users were recruited for each study.

Participants also completed written background questionnaires at the start of each study session. From these, we can describe additional attributes of the actual participants. They ranged in age from roughly 30 years to roughly 60. They ranged in education from "some college" to Masters Degree, with most having a college degree. Subjects were asked a series of attitude questions relating to Web use and personalization. They indicated that they did not read Web privacy statements, but told us additional privacy certifications were important before giving personal information on the Web. Subjects were knowledgeable about browser cookies. They spent considerable time on the Web researching and buying products (about 5 hours per month). Subjects in Studies 2 and 3 visited the HP, Dell, IBM, Compaq, and Amazon sites more than twice a month on average, and visited Sun about once a month on average.

2.4 EXPERIMENTAL PROCEDURE

The experimental procedure—for both group and individual sessions—began with a pre-session questionnaire to collect demographic and job-related information. In group sessions (Study 1 and 2), the experimenter led subjects through three scenarios. Each scenario was presented using a storyboard approach where participants saw screen shots, images of realistic Web pages, and heard how Pat User used different personalization features to complete tasks. This approach is similar to the Design Walkthrough technique for usability evaluation (Vredenburg, Isensee, & Righi, 2001). The experimenter read each task scenario script to participants accompanied by presentation of the computer-based prototype. The first scenario concerned buying a server and a mix of desktop and notebook systems for a new department of ten people who were beginning a new project. The second scenario was about upgrading the server to handle the workload of an additional 10 people and buying additional desktop and notebook systems for them. The third scenario focused on buying accessories – in this case zip drives – for Pat's entire department.

Each scenario presented between several personalization features (e.g., presentation of accessories constrained to only those compatible with a previously purchased machine; presentation of only those servers that match a user's business characteristics as determined by earlier searches). Each scenario presentation was about 20 minutes long and involved presentation of about 10 Web page screen shots.

In Studies 1 and 2, following the presentation of each scenario, the experimenter facilitated a five to ten minute discussion with the participants, covering the features presented. Comments were recorded on paper flipcharts in the room by a second experimenter. These notes were saved for later analysis by the research team. Participants also completed post-scenario questionnaires in which they provided their ratings for each personalization technique covered. After the third and last scenario, participants filled out the post-session questionnaire. Subjects were then debriefed before receiving payment. During the conversational debriefing, important comments were again recorded by the second experimenter.

In individual user sessions (Study 3), six task scenarios were provided in written form to the participant, who completed the tasks using an interactive personalized prototype. The order of the presentation of tasks was counterbalanced across subjects using a Latin Squares design. Each scenario now included between 3 and 4 features. Participants read the scenario description and then attempted to complete the task described using a prototype system described below. In the individual sessions, participants were encouraged to "think aloud" as they completed tasks. Participants' think aloud protocols provided a wealth of data, particularly regarding subjects' emotional reactions. Think alouds also provided subjects the opportunity to explain the rationale behind their decisions and actions. A second experimenter recorded think aloud comments and other commentary spoken by participants, observations of actions and behavior of subjects while using the interactive prototype, interesting user errors, and so on.

As in the earlier studies, subjects filled out questionnaires asking about their reactions to the features presented in each scenario and soliciting comments on the prototype site's design. Participants also completed a post-session questionnaire form and were also asked

how a personalized, adaptive site might influence their future interactions and transactions with the site.

2.4.1 User Study 1

We reviewed the feature list and selected features for inclusion in Study 1 and, as stated earlier, ensured that at least one feature from each of the 14 categories listed in Table 1 was included. We selected features for each scenario by clustering them as appropriate so their use could be illustrated in the context of complete tasks. We established a specific number of scenarios for presentation (three in this case) so that we would be able to present the scenario, provide some illustrations of personalization features in use, have a small group discussion, and allow participants to fill out a questionnaire about the features before the next scenario. In this study, each scenario included between 6 and 9 features. Three scenarios and the post-session questionnaire fit within the two-hour session time frame.

For Study 1 there were a total of 5 two-hour sessions with 3 to 5 participants in each, for a total of 20 participants. The study was carried out in a usability lab specifically set up for a group session.

2.4.2 User Study 2

For User Study 2, we reviewed the results of Study 1 and made several adjustments. We made changes to the personalization features included in the study, adding 4 features and modifying existing features so that we could obtain more specific user reactions. We then modified the scenario and storyboard presentations accordingly. The research team also modified the post-scenario questionnaires so as to change the ranking of features within scenario to individual ratings on a 1 to 7 scale. In this study, each scenario included between 5 and 13 features.

Study 2 was conducted over a four day period in October of 2001. There were a total of 5 two-hour group sessions with 2 to 6 participants in each one, for a total of 23 participants.

2.4.3 User Study 3

Feedback from User Study 1 and 2 was incorporated into the design of User Study 3. The subject ratings from the first two studies revealed the best-received personalization features: 18 features remained above a clear threshold level, and Study 3 tested only these top 18 features. One existing feature was split into two separate ones so that we could obtain

more focused user reactions to its component features. This study involved individual participants recruited from the external target customer population and these individuals controlled a hands-on interactive mid-level fidelity prototype to carry out six typical tasks. Results of the study provided the researchers with more in-depth design information about the most highly rated personalization features and the impact of these features on site visitation and purchases.

Study 3 was conducted during two weeks in December, 2001. There were a total of 22 two-hour individual user sessions.

3 RESULTS AND DISCUSSION

Our studies provided many fine-grained results concerning personalization techniques and facilities (see Karat et al., 2002). In this section, we present results of greatest interest and importance to researchers and developers of user-adaptive systems.

The results reported here include proclivities and impressions gathered from the studies' qualitative data and from participation in discussions with participants during experimental study sessions. Not *every* participant registered negative reactions to, say, the practice of using implicitly gathered information to guide adaptive presentation. But a substantial subset did, and did so compellingly. We believe that negative views of this nature (as well as participants' positive and mixed reactions) must be considered when designing and implementing adaptive Web sites.

3.1 USER CONTROL OF PERSONAL INFORMATION

The clearest result to emerge from our studies is users' fervent desire to be *in control*. This is quite a general result, but was so prominent that it bears mention here. First, this result emerged most strikingly in terms of being in control of one's profile, the persistently stored personal information collected by the system about the user. While this result is certainly no surprise, it is prevalent among all users and is a strongly emotional issue. As such, it can be considered a "deal breaker"—users will not provide personal information and will not perform transactions on a Web site that in their opinion cannot be fully trusted to use profile information in an ethical and private manner.

Users want to be able to review, modify, delete, add personal information at any time. This feature received the highest quantitative rating overall. In Study 2, the feature

"You control all the data kept in your profile and can review and edit it at any time" was rated 1.7 on the 1 (most valued) to 7 (least valued) scale and received an average rating of 1.6 in Study 3. Users also want control over who sees and uses this personal information and whether personal data will be sold or otherwise provided to third parties.

Participants told us they do not want any information that has been collected *implicitly* by the system to become a part of their persistent personal profile. "Implicitly" gathered information implies user interests, user goals, or any other data regarding the user that is inferred by the system based on, for example, navigation paths the user has traveled. Our participants want, instead, only information *explicitly* provided by them to be remembered by the system: "I want the stored information to be based on what I told you, not what you think I've said." Furthermore, if and when there is a question of whether particular data should be stored in the user's persistent profile, participants want control in the form of explicit granting or denial of permission to save these data (as an example from our prototype, users control whether the search criteria provided to the "Help me find what I need" wizard should be remembered in their profiles; see Figure 3).

3.2 USER CONTROL OF CONTENT

In addition to having control over their profiles, participants expressed the more general desire to have more control of the overall interaction and the information they are shown on a site. This result is perhaps at odds with the goals of an adaptive site: as pointed out by Fischer (2001), an adaptive system implies some loss of user control because the system is making content or navigational decisions. Participants were happy with adaptive content based on information explicitly provided by them but mixed in their reactions to adaptivity based on implicit information about them or their previous behavior on the site.

3.2.1 Adapting content based on explicit and inspectable information

To be more specific, participants liked content-filtering and -refinement based on information explicitly provided by themselves. For example, in the "Purchasing a server" scenario, participants used the "Help me find what I need" portion of the PersonalBook, supplied information regarding the price range, intended functionality, and/or number of supported clients for a desired server machine (see Figure 2), and in response the site displayed a page containing a filtered view of all servers, showing only those servers that match the user's requirements. Users found this "help me find" wizard to be an extremely

useful feature and highly rated the generalized notion of content filtering based on user supplied information as exemplified by the wizard (e.g., the average rating for this idea was 1.9 in Study 3).

Participants also appreciated when content was filtered based on openly inspectable personal profile information—for example, the list of products owned by the user and readily viewed in the user's PersonalBook. In the "Compatible memory" scenario, participants opened the "My Products" portion of the PersonalBook, selected a specific computer, thereby navigating to the compatible-accessories page, then found memory compatible with the selected computer. Here, the site filters the list of all memory chip sets to just those compatible with a specific computer. Being able to thereby purchase memory *known* to be compatible with the user's system was perceived as a great benefit. This feature was popular enough to be included among the top 18 features that were included in Study 3—in Study 2 it was rated an average of 2.4 on the 1-to-7 scale—and in Study 3 it received a rating of 2.5.

3.2.2 Inferring what the user "wants" to see

Contrasting with the above results, adaptive content based on *implicit* information, such as previous navigation paths, met with mixed reactions. An example of adaptive content based on previous navigation can be seen in the "Compatible memory" scenario. There, participants selected the ThinkPad system listed in "Products That I Own," thereby navigating to that computer's compatible-accessories page, and then *asked* the site to show compatible memory chip sets. Later, participants returned to "Products That I Own" and selected the desktop computer listed therein; this time, when the compatible-accessories page was shown, it automatically displayed memory options with no user-initiated action required. Here, the site is making the "assumption" that because the user had previously looked at memory compatible with the ThinkPad he might be looking for memory compatible with the subsequently selected desktop computer—that is, the content is being adapted based on a previous navigation path. A number of subjects expressed the view that this particular scenario could be interpreted as "pushy" or irritating on an eCommerce site: some thought of it as a marketing gimmick, intended to sell more memory.

The *idea* of basing content on previous navigation was rated positively by participants: while in Study 1 this feature was only ranked fourth (average 4.1) out of the six features shown in the same scenario, the average numeric rating on questionnaires in Study 2

was 2.4 and in Study 3 was 2.8. In our prototype, the demonstration of adaptive content based on an inferred plan worked perfectly and we believe this, at least in part, accounts for the relatively high quantitative ratings for this feature. It is also plausible that, in Study 2, participants were better able to appreciate this feature after the experiment facilitator explained *the rationale* for the automatically generated content.

At any rate, the qualitative results for this feature are particularly interesting. Some participants were adamant in their disapproval of this feature, and their opinions were so clear and passionately articulated that they must be considered by developers of adaptive eCommerce sites. One participant stated the explicit-implicit problem succinctly: "I like to have more control and less assumptions made." Other participants expressed skepticism that this feature could be implemented in a way that would make it generally valuable to them. Specifically, participants in all three studies expressed some level of concern about a system's ability to do this *well enough* to be useful. A number of subjects declared they would prefer that the site not attempt to infer their intentions, declaring their opinion that "computers are not smart enough" to do a good job of inferring users' goals and plans. Some expected that trying to infer a user's intentions would only lead to an annoyingly poor result, and they cited the Microsoft® Office Assistant (also known as "Clippy," the paperclip) as exemplary support for this opinion. So the message was this: "adapting content based on past navigation *would be* a nice thing, but we don't believe you can do it well, so don't do it at all."

This finding is certainly relevant for eCommerce users, who may interpret adapted content as a marketing maneuver, but seems to be a more generally applicable finding as well. Refining this result further, adapting content based on *immediate* context was met with favor by our participants. For example, in the "Purchasing a Server" scenario the content of the "Help me find what I need" wizard (the central question and the user-response fields) is based on which Web page is currently visible. This is the simplest possible case of content adaptation based on implicit navigation information, in this case using knowledge of the page the user is visiting when she opens the "Help me find what I need" search tool. From this result, it appears that using the immediate or *current* context to influence concurrent or immediately ensuing content is seen to be useful, whereas attempting to infer *current* goals based on navigation or other information from the *past* was not universally welcome.

Participants mentioned "shifting goals" as a problem that they believe would not be handled well by systems that use inferred goals to guide current and future presentations. They realized their needs, even on a single site, change over time, and they do not want adaptive behavior of the site to be based on their obsolete goals. Overall, participants agreed that there ought to be a logical "limit" regarding how far back into one's past history the system should look for implicit information about goals and interests, and further that this limit should be the current "session." That is, the consensus view was that a user's past history should have a limited *life span*: it is acceptable for the system to adapt content based on the user's current task and navigation context and even the user's context and history since she logged on today, but it should not look at past behavior beyond this point to a disjoint time or session in the past.

A further complicating issue relating to the implicit gathering of information about a user's past interaction history is: Is *all* navigation behavior and clickstream data relevant or meaningful? And therefore exactly when should such information *not* be collected by the site? Participants also expressed the desire to be "invisible" during exploratory sessions on a site; they did not want data collected implicitly about them and tied to their profile while they are in this exploratory and information-seeking role unless they specifically authorize it. As one participant explained, "Hey, I might just be knocking around the site for a while...it doesn't mean anything, and you'll fill up my profile with a lot of junk if you implicitly collect that information. Wait until I know that I'm really after something before you start collecting data about what I'm doing, and let me tell you when that is." Thus sites that base adaptation on users' navigation history ought to offer multiple user *roles* or *persona*, including "anonymous" wherein no data about the user is collected.

3.3 ADAPTATION BASED ON TRANSIENT INFORMATION

In one of the scenarios described above, the prototype demonstrated content adaptation based on automatically identified temporary information. In particular, the scenario story depicts the user dialing in from home. The prototype site detects a slow connection speed and as a result, images and animations are not initially downloaded and displayed; in place of each image is a textual link that may be clicked to download and view the image files for the current Web page (see Figure 4). The idea here is that the site is adapting the content (to optimize download time) based on the connection.

Based on the results discussed in preceding sections, we can assume users do not want to be "surprised" in the sense of "where did that content come from?" Users seem to have similar expectations of a dialog with an application or Web site as with other people. That is, it appears that users desire responses from a Web site to follow logical conversational conventions (Grice, 1975; Searle, 1969); for example, they expect system responses to be relevant to the ongoing dialog and to be as informative as required but not more so. This result coincides with the experimental work of Reeves and Nass (1999). After conducting numerous studies, Reeves and Nass concluded that users' expectations regarding their interactions with computers and other media are based in large part on social interactions in real life, that is, users expect computers to obey rules that come from the world of interpersonal interaction. When an application provides more information—or, in some manner, other information—than what is expected, the user may not immediately know what to make of the "extra" information. In discussing many features, we heard from our users that they want to *understand*, without difficulty, why and how the computer side of the ongoing bilateral conversation, represented by the content the Web site displays, chose to "say" what it has displayed.

So it would appear that users might not highly rate the automatic adaptation of content based on background information such as the current line connection speed. And in fact user response to this feature was not enthusiastically favorable—although it was rated higher than neutral. After Study 2, content adaptation based on such transient information as line speed *was* rated sufficiently high to be among the top 18 features (and thereby included in Study 3): it received a rating of 2.5 in Study 2. In Study 3 it scored 3.5, where 4 is neutral. The difference between studies may be reasonably understood. In Study 2, the experimenter drew attention to this feature as he or she led a tour of the site. In Study 3, in which users directly interacted with the prototype, this feature had to be "discovered" by the user; participants were focused on accomplishing their eCommerce tasks and therefore may not have even noticed that images were absent and textual links appeared in their places.

More importantly, the fact that some participants expressed the opinion that "adaptability based on temporary information is good" makes sense when we consider that in the second study the facilitator not only pointed out how this feature was manifested in the Web pages being viewed but also *explained why*. Even in Study 3, the prototype site

explained the reason for its adaptive behavior in the text of the hyperlinks that appeared in place of images—this text indicated that the site detected a slow connection speed (see Figure 4)—so if participants discovered this text they would understand what the site was doing and why.

Other types of transient information that can factor into content decisions include the user's behavioral state (Syeda-Mahmood, 2001), the user's affective state (Hudlicka & McNeese, 2002), the client device type (e.g., PDA, cell phone, PC), and the user's Web browser and its capabilities (e.g., dynamic HTML (DHTML) features vary by browser, JavaScript is/is not enabled) (see also Kobsa et al., 2001). A user should not be surprised to find that the content of a site's Web pages when viewed on a handheld device is different than pages displayed when viewing the site on a PC's larger screen. However, from our results we can conclude that when transient information is a factor governing content decisions, the resultant content's rationale should be made clear to users.

3.4 COLLABORATIVE FILTERING

When a purchase transaction is about to be completed, many existing eCommerce sites recommend additional "related" products to the user. Here, "related" can imply many things, but in our study we tested a specific class of recommendation: "People who bought the specific product you are now purchasing also bought products *X*, *Y*, and *Z*." In particular, when a laptop was being purchased on our prototype site, the additional products suggested to the buyer included a carrying case, extra batteries, and a power adapter for cars and airlines. This is a simple form of *collaborative filtering* or *recommender* technologies (e.g., Resnick & Varian, 1997; Schafer, Konstan, & Riedl, 2001; Burke, 1999). When applied in the context of eCommerce, these technologies use the buying behavior of prior customers to attempt to "predict" which products a new customer may be interested in purchasing.

Our participants had mixed reactions to this feature and overall did not rate it very highly relative to other features: it was rated 3.2 in Study 2, a low enough rating that it was not included in Study 3. Perhaps one reason is what one participant stated: "I am not like other people, I have different needs." Or, this feature might be perceived simply as an intrusive marketing ploy. As another participant declared, "[This feature] may be helpful, may be obnoxious." Looking a bit deeper, it appears that user attitudes about collaborative filtering may in fact be influenced by a variety of factors including the type of eCommerce

site involved, the particular product being purchased, the type of user, and the reason an item is being purchased. Our studies involved a specific class of user, a particular type of e-business site showcasing a particular category of products and services, and a specific type of purchasing scenario wherein users were business managers buying computer equipment for their companies. They had direct, objective needs, driven by our instructions in the study scenarios and driven by specific business requirements in their actual jobs. What other customers purchased is not interesting or important to these users.

The upshot is that our users had equivocal reactions to a feature that is nonetheless in extensive use on eCommerce sites. Perhaps, if our participants were purchasing books or music CDs, they would indeed be happy to see the recommendations of other, similar customers. The clear conclusion, however, is that individual personalization strategies, such as recommender systems, are not necessarily effective or beneficial across the spectrum of eCommerce activities; instead, user opinions and desires regarding such technologies may depend on multiple dimensions of the eCommerce scenario in question, such as the type of product being purchased, whether the purchase is for one's self or for one's company, and whether subjective parameters such as taste or genre are relevant to the item being purchased.

4 CONCLUSION

Design of personalization or user-adaptive systems (or any software technology) cannot occur "in a vacuum," specifically, it cannot usefully proceed without assessing the value and usefulness to users of the concepts proposed and implemented by researchers and developers. Empirical studies of actual users help reinforce, contradict, and refine designs to better accommodate and satisfy users in accomplishing their tasks. Though our studies involved a specific class of user, a distinct type of eCommerce Web site (selling a specific category of products and services) and specific purchasing scenarios, we feel that our results regarding users' attitudes towards personalization and adaptive systems are more generally applicable.

One important result of this work concerns the explicit-implicit question, the collection and use of explicit or obvious information versus implicit information about the user. It appears that users are happy to have the *current* (and thereby obvious) context and

even context from the current session (*recent* past) influence the presentation of new content. Along similar lines, content filtering based on information or criteria provided *explicitly* by the user, or explicitly selected by the user, or that is explicitly accessible in the user's profile was a very highly rated feature. However, some participants were skeptical of systems that try to predict a user's current goals or needs and adapt content accordingly based on implicit information from some disjoint time in the past. Interestingly, users reacted favorably to such a feature *in theory*, but felt that systems would not, or cannot, do this well *in practice*. This view was based in part on their experiences with existing products that attempt to infer user goals but do so poorly. Thus there is another important lesson here: the rush to market of immature technologies can have a deleterious hysteresis effect, tainting users' later opinions of similar products which may in fact work well.

In addition, the participants in our study were not strongly in favor of collaborative filtering techniques as exemplified by additional product recommendations made at purchase time ("people who purchased this product also bought …"). Our users were loathe to have their *own* behavior from previous sessions govern what they see on a site—it is therefore logically consistent for them to feel likewise about the use of the prior behavior of *other* customers to infer or project the content the user is interested in viewing. Remember, however, that our studies involved products with particular characteristics, namely computing machinery and software for an enterprise. On the other hand, it appears that customers purchasing items such as books and music are happy to see recommendations from other customers regarding books and music of a similar genre. The apparent upshot is, not all personalization and adaptation technologies can necessarily be applied in all eCommerce situations; careful assessment of user needs and desires given the specific eCommerce domain and scenario must be performed before deploying such technologies.

Our findings also support the fact that, overall, users want to feel as if they are in control. A significant portion of a user's sense of control is whether she can readily *make sense* of the interaction with a site. Not understanding why the site is displaying particular content engenders for users a sense of loss of control ("what's going on here?"). Conversely, understanding the how and why of a site's content choices and behavior grants users more of a sense of being in command. When content is based on explicit user requests, users obviously understand the causal relationships that give rise to the content. When immediate

context is used to guide content, users can easily infer the source of the site's content. When past behavior or other non-obvious implicit information are used to generate content, the users' disconcerted reaction—"where did this come from?"— tells users they are not in control. Content whose origin is obvious or readily inferable is met with favor. Content that is based on something the user, or other users, did at some temporally distant time is often met with disapproval. Finally, personalized Web sites must provide users the feeling and the reality of control over the information contained in, and the use of, their personal profiles.

REFERENCES

Ardissono, L. & Goy, A. (1999). Tailoring the interaction with users in electronic shops. In *Proceedings of the 7th International User Modeling Conference (UM'99)* (pp. 35-44).

Burke, R. (1999). Integrating Knowledge-Based and Collaborative-Filtering. In *Proceedings of AAAI 1999 Workshop on AI and Electronic Commerce* (pp. 14-20).

Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, *11*, 87-110.

Chin, D.N. (2001). Empirical evaluation of user models and user-adapted systems. *User Modeling and User-Adapted Interaction*, *11*, 181-194.

Fischer, G. (2001). User modeling in human-computer interaction. *User Modeling and User-Adapted Interaction*, *11*, 65-86.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41-58). NY: Academic.

Hudlicka, E. & McNeese, M.D. (2002). Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task. *User Modeling and User-Adapted Interaction*, *12* (1): 1-47.

Joerding, T. (1999). Temporary user modeling for adaptive product presentation in the Web. In *Proceedings of the 7th International User Modeling Conference (UM'99)* (pp. 333-334).

Karat, J., Karat, C-M., Brodie, C., Alpert, S. R., & Vergo, J. (2002). *Personalizing Interaction: Customer and Business Value of Personalization Features in e-Commerce*. Submitted.

Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalized hypermedia presentation techniques for improving online customer relationships, *The Knowledge Engineering Review* 16(2), 111-155. http://www.ics.uci.edu/~kobsa/papers/2001-KER-kobsa.pdf.

Kramer, J., Noronha, S., & Vergo, J. (2000). A User-Centered Design approach to personalization. *Communications of the ACM*, *43*(8), 45-48.

Milosavljevic, M. & Oberlander, J. (1998). Dynamic Hypertext Catalogues: Helping Users to Help Themselves. In *Proceedings of the 9th ACM Conference on Hypertext and Hypermedia (HT'98)*. Also at http://www.dynamicmultimedia.com.au/papers/ht98/.

Paris, C. L. (1987). *The use of explicit user models in text generation: Tailoring to a user's level of expertise*. PhD Thesis, Columbia University.

Paris, C.L. (1993). *User Modeling in Text Generation*. London: Pinter.

Reeves, B. & Nass. C. (1999). *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. Stanford, CA: CSLI Publications/Cambridge University Press.

Resnick, P. and Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3) 56-58.

Rich, E. (1979). *Building and Exploiting User Models*. PhD Thesis, Carnegie-Mellon University.

Schafer, J.B., Konstan, J., & Riedl, J. (2001). E-Commerce Recommendation Applications, *Journal of Data Mining and Knowledge Discovery, 5*(1/2), 115-152.

Searle, J.R. (1969). *Speech acts: an essay in the philosophy of language*. London: Cambridge University Press.

Sleeman, D. & Brown, J.S. (1982). *Intelligent Tutoring Systems*. London: Academic.

Syeda-Mahmood, T. (2001). Learning and Tracking Browsing Behavior of Users using Hidden Markov Models, In *Proceedings of Make IT Easy 2001, The IBM Ease of Use Conference*, http://eou2.austin.ibm.com/easy/eou_int.nsf/EasyPrint/1859.

Vredenburg, K., Isensee, S., & Righi, C. (2001). *User Centered Design: An Integrated Approach***.** Upper Saddle River, NJ: Prentice Hall.

FIGURES

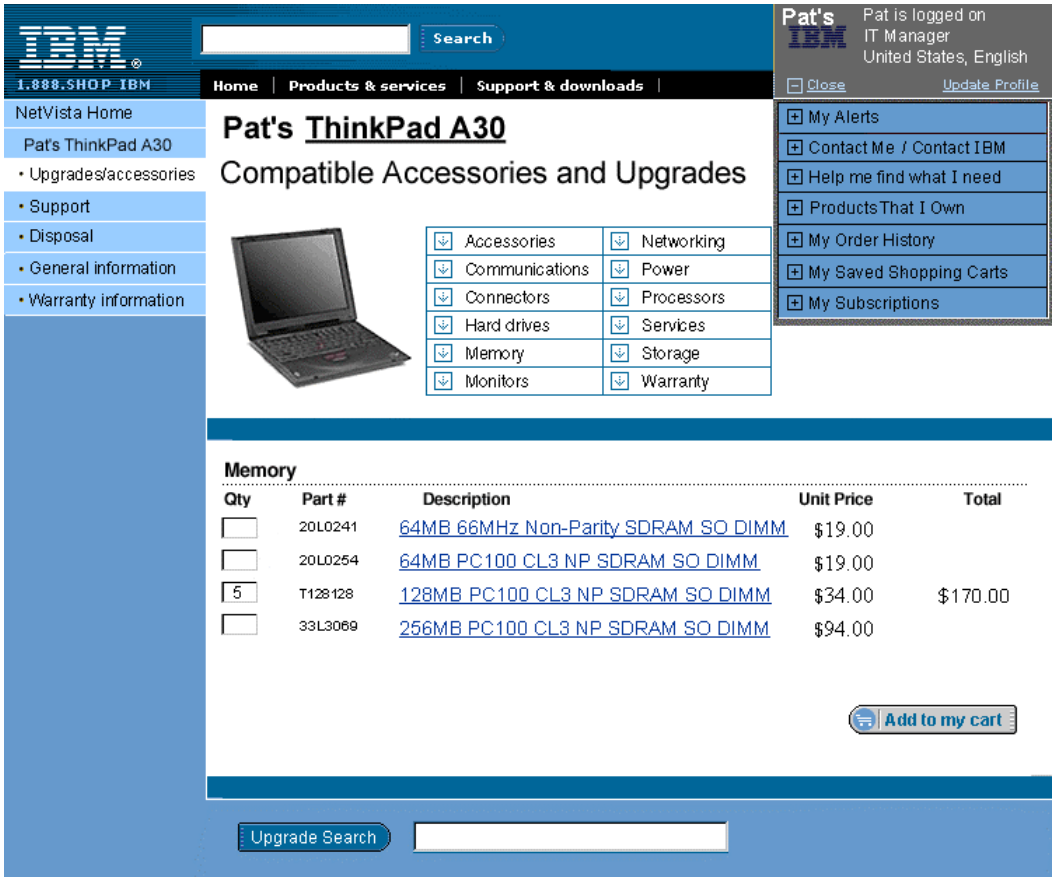

Figure 1. A page from the prototype Web site used in the studies. In the upper right is the personalization portion of the site masthead, showing that a user whose ID is "Pat" is logged on. The user may edit her personal profile information at any time by clicking the "Update Profile" link. On the right below the masthead is the PersonalBook, which has been opened by the user and overlays a portion of the main Web page. Individual sections of the PersonalBook are viewed by clicking on the corresponding plus-sign or label. The main Web page here shows accessories that are compatible with a specific computer.

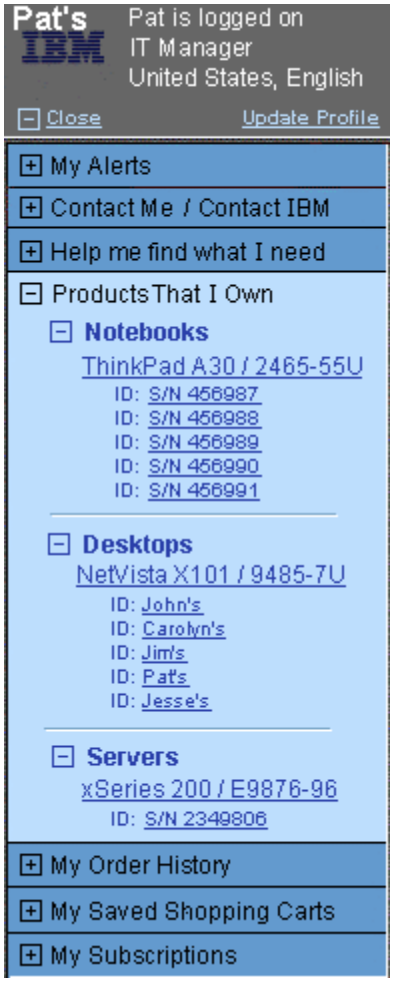

Figure 2. The PersonalBook with the "Products That I Own" section open. A user may click on any product listed here to navigate to the page showing only those accessories and upgrades that are compatible with that specific product—see the main Web page shown in Figure 1.

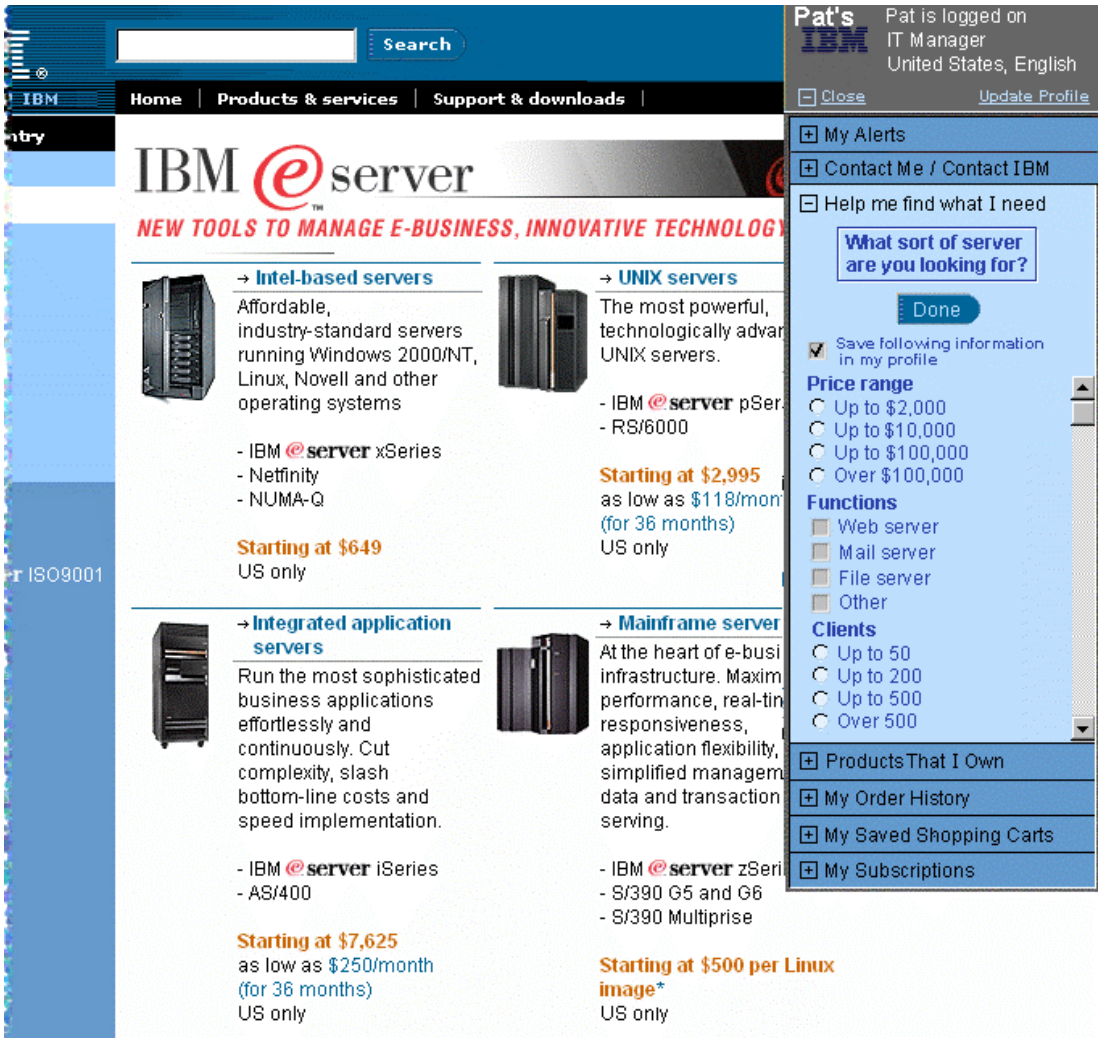

Figure 3. The "Help me find what I need" portion of the PersonalBook. This tool acts as an assistant to the user in locating particular products or services on the site. The specific question posed to the user and the user-reply fields are context-dependent.

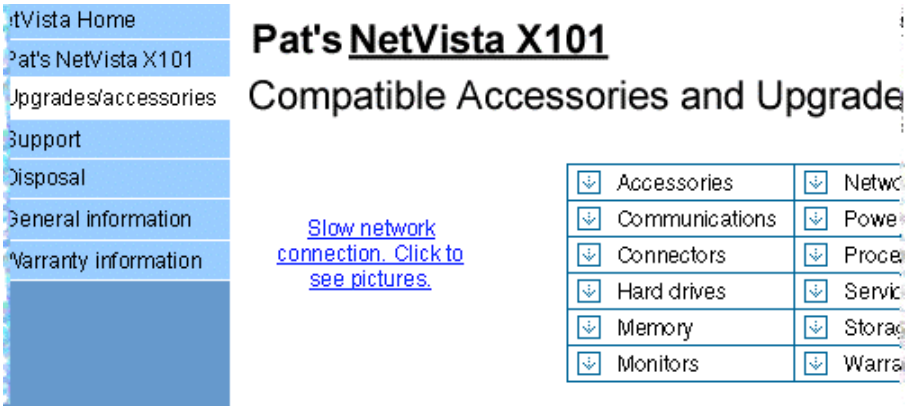

Figure 4. When a slow connection speed is detected, the prototype site does not automatically download image files and in place of images displays the text link shown here. A user can view all pictures on the current page by clicking the link.

Authors' vitae

Sherman R. Alpert
IBM T.J. Watson Research Center, PO Box 704, Yorktown Heights, NY 10598 USA
(Address is the same for all authors)

Sherman Alpert has been at the IBM T.J. Watson Research Center since 1987. He received a B.S. in Computer Science from the State University of New York at Stony Brook and M.A. in Computing in Education from Columbia University's Teachers College where he has pursued additional studies. He has been involved in research and software development in a variety of domains including educational technology, human-computer interaction, multimedia, and object-oriented programming and design. His current work involves user experience research. He serves on the Editorial Boards and review boards of several journals and on several conference and program committees.

Dr. John Karat:
John Karat is a Research Staff Member at IBM's T.J. Watson Research Center since 1987. He received his Ph.D. in Cognitive Psychology from the University of Colorado in 1982. He is chairman of the IFIP Technical Committee on Human-Computer Interaction (IFIP TC13) and is currently working in the areas of Privacy and Search in interactive systems.

Dr. Clare-Marie Karat
Dr. Karat is a Research Staff Member at the IBM T.J. Watson Research Center. She received her B.A. in Psychology with Honors from Stanford University, and a Ph.D. in Social Psychology from the University of Colorado at Boulder. Dr. Karat conducts human computer interaction (HCI) research in the areas of privacy, personalization, and conversational interface technologies. Dr. Karat is an international expert in cost-justifying human factors work and develops innovative user interface designs and methodologies. In 1998, Dr. Karat wrote the Computer User's Bill of Rights which was featured in numerous publications around the world, is part of IBM's Ease of Use instructional materials, and is featured on the ibm.com web site. Dr. Karat led the development of the personalization strategy for ibm.com and the results reported here are part of that research project.

Dr. Carolyn Brodie
Dr. Brodie is currently a Research Staff Member at IBM's T. J. Watson Research Center where she works in the area of HCI research. She received her B.S. in Computer Engineering from Iowa State University in 1984 and her M.S. and Ph.D. in Computer Science from the University of Illinois at Urbana-Champaign in 1997 and 1999 respectively. Her research interests include human-computer interaction and interactive planning technologies. Most recently her work has concentrated on how privacy and personalization technologies can be made more usable and understandable to both e-commerce companies and their customers.

John G. Vergo
John Vergo is the manager of the User Experience Research group at the IBM T.J. Watson Research Center in Hawthorne New York. His research interests include human-computer interaction, user-centered design methods, multimodal user interfaces, e-commerce user

experiences, speech recognition, natural language understanding. scientific visualization, 3D graphics and software development methods. He has a BS in Mathematics and Psychology from the University at Albany, and an MS in Computer Science from Polytechnic University.