Serving Citizens' Needs:

Minimizing Hurdles to Accessing Government Information Online[*]

Eszter Hargittai[**]
http://www.eszter.com
Sociology Department
Princeton University

This paper is a pre-print copy of the article that appeared in IT & Society available online at
http://www.stanford.edu/group/siqss/itandsociety/v01i03/v01i03a03.pdf

Abstract

With the rapid spread of the Internet across society, government institutions are taking advantage of using digital technology to distribute materials to citizens. Is merely having a Web site enough or are there certain usability considerations to which site creators must adhere in order to assure efficient access to online materials? Here, I report on a project that looked at people's ability to locate various types of content online. In particular, I focus on people's ability to find tax forms on the Web. Findings suggest that people look for content in a myriad of ways and there is considerable variance in how long people take to complete this online task. Users are often confused by the ways in which content is presented to them. In this paper, I discuss two common sources of confusion in users' online experiences with respect to locating tax forms online: 1. URL confusion; and 2. page design layout. In addition to describing these problem areas, I also suggest ways in which these two sources of frustration could easily be curtailed yielding less exasperating and more productive user experiences.

Introduction

There are many studies that look at how people use the Internet (for a review, see DiMaggio et al. 2001) and in particular what types of content they view online (e.g. Howard, Rainie and Jones 2001). There is a separate body of literature that looks at how people use information retrieval systems and, in particular, how people search for information on the Web (for a review of this literature, see Jansen and Pooch 2001). However, these two areas of inquiry rarely intersect leaving the discussion of what people do online in isolation from studies of what people are *able* to do online. Moreover, such literature rarely considers the production side of online content distribution. Does the way in which content is organized and presented online influence people's ability to find their way to material on the Web? In this research report, I present findings from a project that explores people's search strategies in locating content on the Web in conjunction with a look at how Web sites present information to users. I focus on access to government documents, in particular, I explore how users find individual federal tax forms on the Web.

In the first section, I outline the methods used in the project for studying how people locate content online, the sampling methodology, and how the data were analyzed. Then, I report the findings from 97 interviews and in-person observations conducted with a random sample of Internet users. In particular, I focus on how URL confusion and page design layout problems contribute to frustrating experiences for users. Finally, I suggest ways in which these two sources of confusion can be remedied to offer a more productive and satisfying user experience not only for locating tax forms online but other types of government documents as well and, in fact, any other type of material on the Web.

Digital Inequality

Millions of people access the Web daily for financial, health and government information, for job searches, entertainment and numerous other activities (Howard, Rainie and Jones 2001). As daily activities continue to move online, the ability to efficiently navigate the Web for information becomes increasingly important to maintaining a competitive edge and guaranteeing equal opportunity in the various aspects of life including access to government, health and financial services, education, consumer information and a myriad of other realms. As such, Web use skills are

poised to become an important part of people's human capital. Internet skills are a new vital component of human capital in the digital age.

Much attention among both academic researchers (Bucy 2000; Hoffman 1998; Strover 1999) and in policy circles (Benton and Leadership Conference on Civil Rights Education Fund 2002; National Telecommunications and Information Administration 2002) has been paid to what segments of the population have access to the Internet or are Internet users. Access is usually defined as having a network-connected machine in one's home or workplace. Use more specifically refers to people's actual use of the medium beyond merely having access to it. The digital divide is most often conceptualized in binary terms: someone either has access to the medium or does not, someone either uses the Internet or does not. Such an approach to the digital divide has led some to conclude that we can "declare the war won" given that access and use has increased consistently over the past years (NTIA 2002) and the majority of the American population is now online (Compaine 2001).

However, this approach wrongly assumes that gaining access to the Internet obliterates any potential inequality that may result from lack of access to the new medium. There are factors beyond mere connectivity that need to be considered when discussing the potential implications of the Internet for inequality. In addition to relying on basic measures of access to a medium, we need to consider more nuanced measures of use such as user skill. Skill I define as the ability to locate content online effectively and efficiently.

In addition to refining our understanding of what it means to have effective access to a medium (Wilson 2000), it is also important to consider the institutional factors that influence how people use a medium. Individual uses are embedded in a higher-level system organization. Business practices shape the landscape of a new technology's industrial organization, which then affects how it is made available, presented and distributed to users.   In this research report, I discuss findings about online information seeking both at the level of user skill and with respect to organizational practices that influence users' online actions.

Data and Methodology

The study is based on in-person observations and interviews with a random sample of Internet users from a New Jersey county. Respondents were asked to come to a university location for participation. Study participants were given the choice of

using a PC or a Mac both of which were loaded with the three most popular browsing software applications (Internet Explorer, America Online, and Netscape Communicator) to allow respondents to replicate their usual online experience.[1] The computers connected to the Internet on a high speed university network line. Additionally, a program called Don't Panic (Panicware 2001) was used to erase the browser and URL history on each browser program so that each respondent started out with a clean slate and was not influenced by previous users' actions. The search sessions were recorded with a screen capture program that generated audio-visual files of the entire search session.[2]

The researcher sat behind to the left of the respondent and refrained from influencing the respondents' strategies (e.g. never suggested any particular online action, did not answer questions about spelling or whether a certain click would be useful). Respondents were encouraged to look for the information until they found it. No one was cut off from pursuing a search. In some cases when respondents looked frustrated or agitated they were given the option of moving on to another task. However, when a subject simply stated that he or she was unable to perform a certain task, that person was encouraged to try several times before moving on to the next task.

Information about subjects' usual Internet use and history as well as data on demographic background were collected via surveys one of which was orally administered in the beginning of the study session while the other participants filled out online at the end of the study session. The audio-visual files were coded to see whether people successfully completed tasks and how long they took to do so (in seconds). Information about what sites users visited and which online actions led them there was also coded. The audio component of the interviews was transcribed to offer information about users' understandings of their actions and how they felt about their search experience. Hargittai (In press) describes the methodology in more detail.

The findings reported here are based on 97 observation sessions and interviews conducted between the summers of 2001 and 2002 in a New Jersey county.[3]

---

[1] No default page was set on browsers in order not to influence respondents' initial actions once online. The sessions were started off by the researcher asking the respondent to recall – if possible – the default homepage on the computer she uses the most.

[2] The Hypercam program from Hyperionics was used on the PC and SnapZPro program from Ambrosia Software Inc. was used on the iMac. The whole screen was captured, as was every action – e.g. click of the mouse, scrolling – and every comment the respondent made during the search.

[3] The author conducted 77 of the interviews; two research assistants administered the remaining twenty.

Unlike respondents in many related studies (McDonald and Spencer 2000; Wang, Hawk and Tenopir 2000), the participants in this project represent a diverse group of Internet users. They range in age from 18-81, half (51.5%) are women (see Table 1 for details). Participants' occupations range from real-estate agents, environmental policy analysts, blue-collar workers to office assistants, teachers, service employees and medical professionals in addition to students, unemployed and retired persons.

	Mean	St. dev.	Median	Minimum	Maximum
Age	42.96	15.86	42	18	81
Education [a]	N/A	N/A	College	Less than high school	Ph.D.
Family income [ab]	N/A	N/A	$80-89,000	$17,500-19,000	>$250,000
Number of years since first use of the Internet	6.28	3.38	6	0	16
Number of hours browsing the Web weekly	8.62	9.39	7	8 minutes	70 hours

Table 1. Descriptive Statistics About Participant Sample.

(a) Education and Family income have no means as those variables were collected categorically. (b) The average median household income in this county in 2000 was almost $56,613 (based on Census data) and the mode for household income was $75,000-99,999 so this sample is what we may expect for local Internet user demographics.

The group is also diverse regarding Web use frequency and history. Participants' Web use ranges from just a few minutes a week to over 30 hours weekly. The group is similarly diverse in its overall experience with the medium. One person went online the year of the study with an additional 13 percent only having used it for two years or less. However, many – 39 percent – of the subjects had been users for 5-7 years. There are also several long-term users among the respondents with 15 percent having had their first exposure to the Internet over ten years ago.

Searching for tax forms online

Differences in search effectiveness and efficiency

Participants were asked to look for tax forms online, in particular, for the individual federal 1040 form. Respondents used a diverse set of methods to find such government material. When given unlimited amount of time, 93 percent of participants were able to locate such a tax form on the Web. The seven participants

who were unable to find tax forms online tended to be statistically significantly older, spend considerably less time online per week and have started to use the Internet more recently than those who could successfully complete this task.

There is a large variance in the amount of time respondents spent on this task. Users took anywhere from less than half a minute to almost nine minutes to look for tax forms on the Web with the average participant spending two and a half minutes searching for such a form. The analyses controlled for earlier experience with looking for tax information online but having such past experience had no statistically significant effect on successfully completing the task.

Types of search strategies

Users turned to a variety of strategies to begin their quest to solve this task. Sixty percent of users turned to a search engine whereas forty percent tried a specific URL (Uniform Resource Locator or Web address). Among those who used a search engine, twenty-nine percent turned to Google, twenty-four percent used AOL, fifteen percent searched via Yahoo!, sixteen percent used MSN's search engine and the rest used various other search engines and portal sites (Table 2 summarizes these data). The majority of those who turned to Google or Yahoo for their search tended to be quicker in completing this task whereas those who relied on AOL or went to a specific site as their first move tended to take longer in successfully finding tax forms online.

Search Engine	Percentage of those who used a search engine
Google	29
AOL	24
MSN	16
Yahoo	15
Other	16

Table 2. Search engines used for tax query

Among those who tried a specific URL, the majority tried to guess the address of the Web site for the Internal Revenue Service. The guesses included "irs" with the following three extensions: .gov, .com, and .org. The correct address is www.irs.gov yet only 52 percent of those trying a URL used that address. It turns out that using this strategy did not improve one's chances of finding tax forms due to design concerns on the IRS's Web site (these are discussed in the next section). However,

with better usability on the IRS's Web site, knowing to go directly to www.irs.gov could aid in people finding tax forms quicker than those who do not know the correct online address of the agency.

Thirty-three percent of participants tried to go to www.irs.com and fifteen percent entered www.irs.org in the location bar of their browser (see Table 3 for a summary of these figures). In 2001, the .com and .org addresses redirected to the same commercial site. In 2002, they led to two different sites but both remained unaffiliated with the Internal Revenue Service. Some people tried yet other variations of irs, e.g. "irs.dot.gov" or simply "www.irs" without any extensions for specifying the top level domain. Among those trying to access a site directly, a handful of participants went to a financial site with which they were previously familiar. These people soon switched to search engines because they could not find tax forms on these sites.

Strategy	Percentage of users
IRS related URL	30
Search engine	60
Specific site (other than "irs")	10

IRS domain extension	Percentage of users
.gov	52
.com	33
.org	15

Table 3. Search Strategies Used When Looking for Tax Forms Online

Sources of confusion and frustration

Many users expressed considerable amount of frustration with this task despite the fact that tax forms are available online in a myriad of locations and the Internal Revenue Service itself has a comprehensive site with many forms and other information. There were two major sources of frustration for users: confusion stemming from the use of various Web addresses (anything other than "www.irs.gov") and misunderstandings or puzzlements due to the site organization and layout of the IRS's Web site. I discuss both of these points in detail.

URL confusion

As mentioned above, over half of users knew the correct address of the IRS's Web site: "www.irs.gov". However, note that three of the people who knew to go to "irs.gov" did not enter "www" as the beginning of the address. To this day, one encounters an error when attempting to access the IRS without the "www" in the address. (See Figure 1 for a screen shot of what happens when "irs.gov" is entered in Internet Explorer. All screen shots in this paper are taken from study session recordings.)
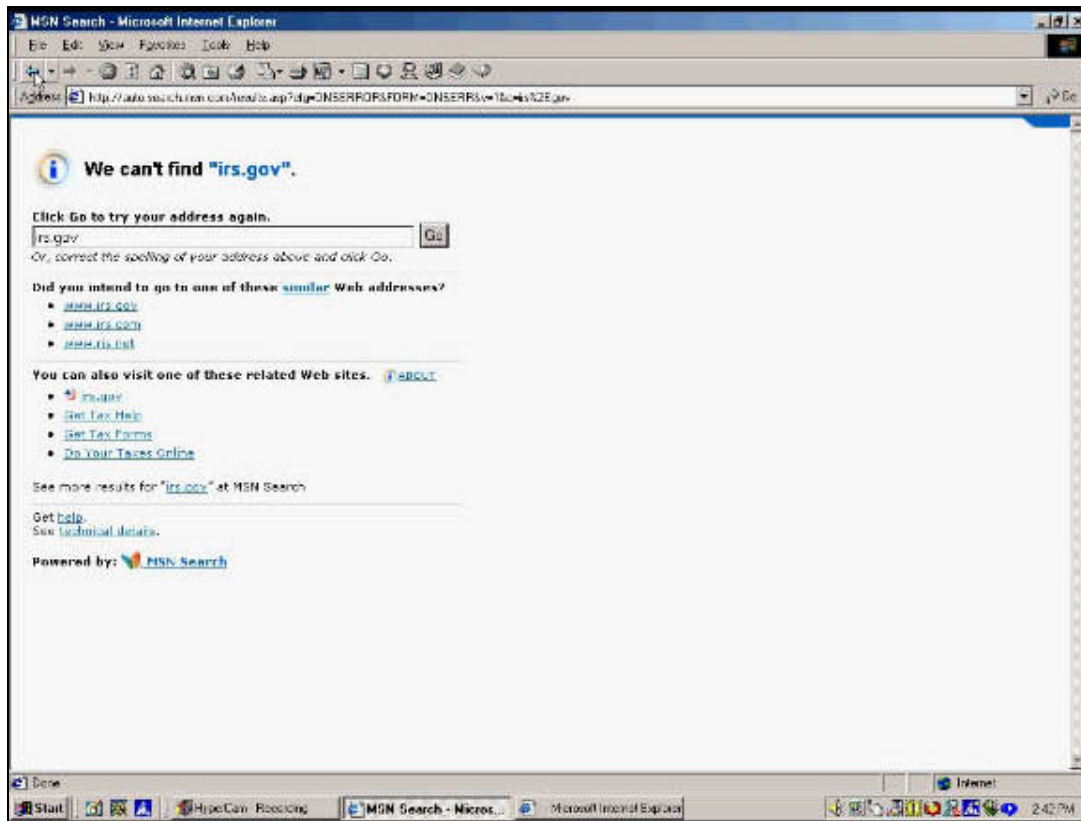

Figure 1. Screenshot of error after entering "irs.gov" in the location bar of Internet Explorer (Fall, 2001)

A more common source of confusion for users was the busy commercial site that comes up under www.irs.com. Until recently, www.irs.org also redirected to this page. The site (see Figure 2 for a screen shot) features numerous links but none of them a direct link to actual forms. Although there is a link called "Individual Tax Forms" it leads to the results page of a search engine with lots more confusing commercial links (see Figure 3). The following were the comments of a 36 year old woman who works for a local government agency during her search process:

[..] Oh ! dot com, no dot gov…wow look at that…that's fancy…so… "get your free"…what is this? [laughs] this is so silly! [laughs]…I mean is this stupid or what? Oh here "department of the treasury", let me just…oh my gosh! I don't know! […] Ok, wait…oh my gosh…here… www.irs.gov…ok, so you get to this ridiculous [laughs] it's not funny [laughs] "tax stats, tax"…oh maybe "electronic services"?…well this is to find, to do it online? Oh just to find forms? So watch, so I clicked on that…oh ok…that didn't come up though… "enter another government", "you will leave the"…oh! "enter another government Web site created, operated, and maintained by that agency"… "continue"…[15 second pause] well that's not very helpful…ok, so I would probably go to […] I would probably just go back to a search engine and type in…[laughs]… "free 1040 tax forms"…free? [excited] Of course they're free…you have to pay for tax forms? C'mon…I get all excited when I see the word "free", and then I'm like "what the heck?" [laughing] hello!… "1040 easy", "tax return online"…"quick tax" see I guess I would search for something that had "dot gov" behind it… "file manager for forms" maybe… "forms and publications" but is that…see that's…oh! Ok that only took about 20 minutes! [laughs]

In reality, this entire search took 2:49 minutes, but the respondent's comment at the end suggests that she perceived the search process to have taken longer than expected given the task. As we can see from her comments – as she narrates most of what she does – she is quite confused about many of the steps in the process. The various commercial sites that came up as she entered www.irs.org did not lead her to the forms for which she was looking. Numerous other people ran into this problem of URL confusion and spent several minutes clicking through from one irrelevant site to another.
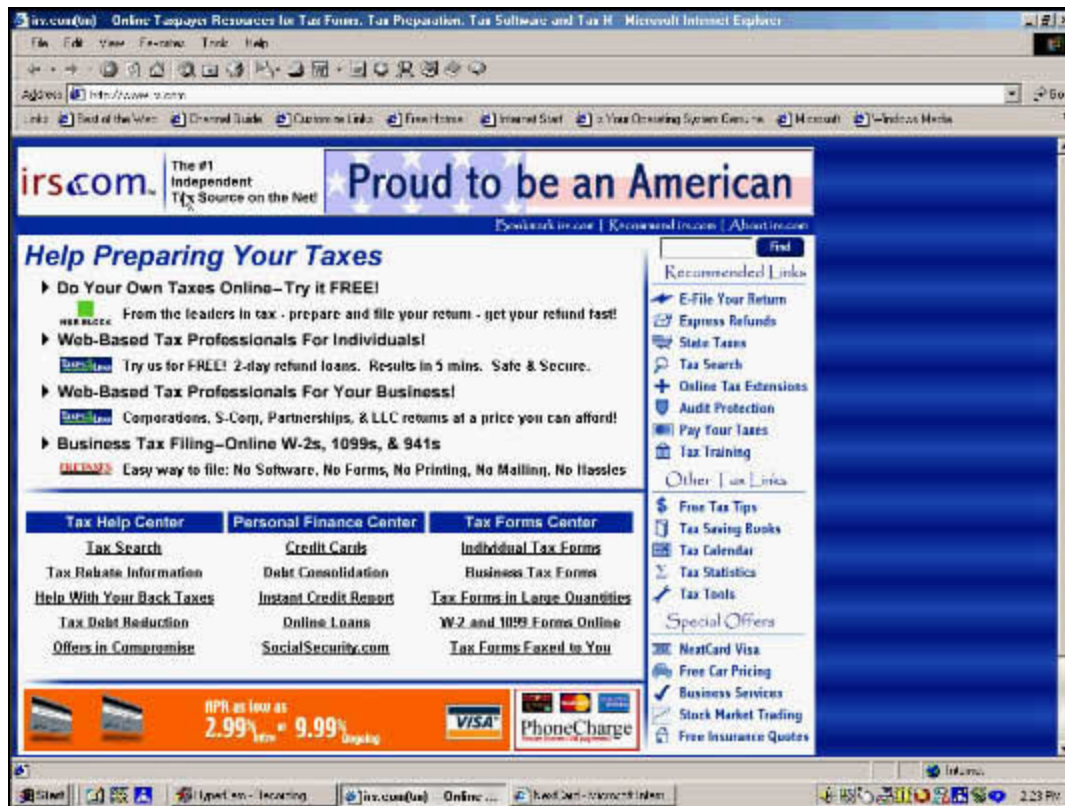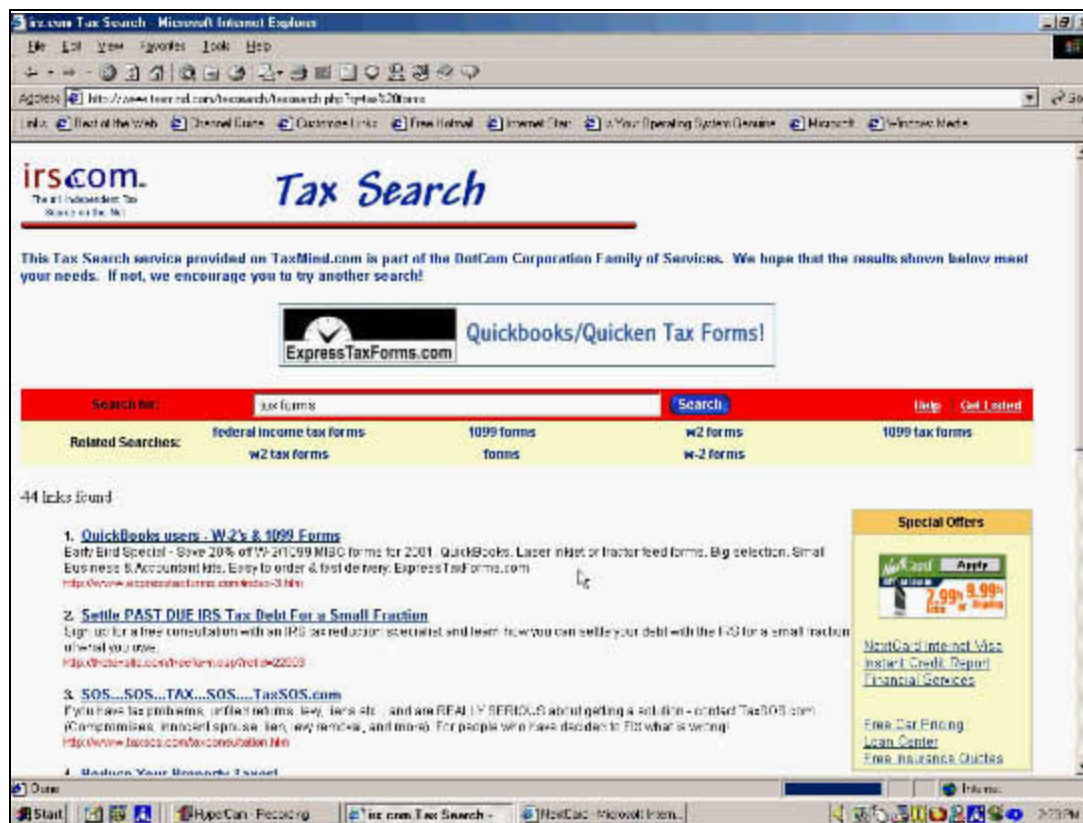
Figure 2. Screenshot of www.irs.com (Fall, 2001)

Figure 3. Screenshot of site linked from www.irs.com (Summer, 2001)

Web site organization and design problems

The effect of design on users is well summed up in one user's reaction to the IRS's 2001 site: "I would go to www.irs.gov [she types it into the location bar and the IRS's site comes up], I believe that is it.. **And that is not it**." [emphasis mine]. This user, a 25 year old college educated woman who works as a financial analyst, was so confused by the design on the IRS's welcome page that she immediately hit the Back button on her browser and proceeded to use Yahoo's search engine instead (see Figure 4 for a copy of the screen this user saw when she went to www.irs.gov). Another user – a 40 year old man who works on medical equipment repair and installation – had a similar impression: "it looks almost [..] comical […] to me it doesn't look like a professional site, I don't know…[laughing] I wouldn't go there to file income taxes or to get a tax form…[laughing] would you?"

Figure 4. Screenshot of www.irs.gov (Summer, 2001)

Those users who recognized that they were on the IRS's site were confused at a later stage in their search. The quickest path to tax forms on the 2001 version of the IRS site was to click on Forms and Publications and then follow a link called Forms and Instructions. Unfortunately, the "Forms and Pubs" link was only available on the bottom of the IRS's home page well below what a user would be able to see when the

page initially came up in her browser. Not everyone saw this link and so these people clicked on something else that could possibly lead to forms on the Web site, but this involved a guessing game. Such ambiguity is frustrating for users. As one male participant – a 69 year old retired marketer of scientific equipment – commented with irony in his voice: "this is crazy, your government working for you" after frustratingly scrolling up and down on a page that seemed like it may have forms but didn't (see Figure 5). In the end, he did notice the "Forms and Pubs" link on the bottom of the page, but only after getting somewhat frustrated with this experience.
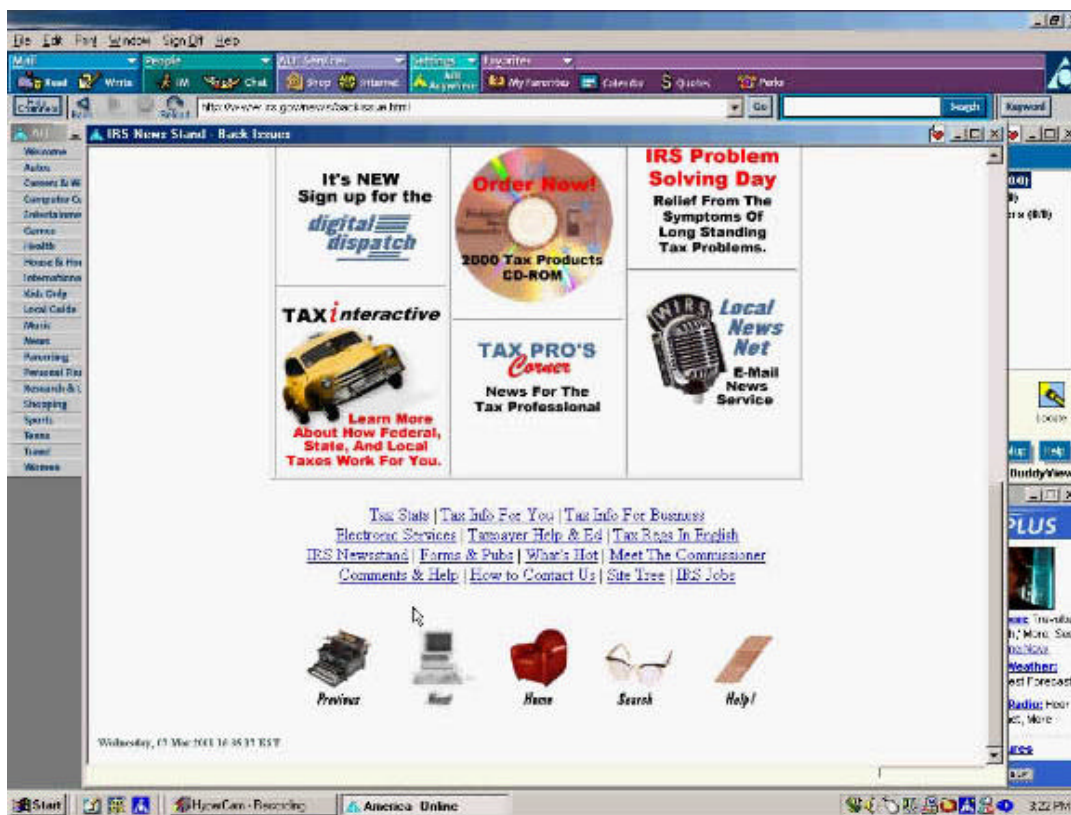

Figure 5. Screenshot of an IRS page (Fall, 2001)

Even those who quickly found their way to the Forms and Instructions page were in for a surprise: no obviously visible link to forms. Figure 6 shows the part of the Forms and Publications page on the 2001 IRS site that was viewable on a standard 17" monitor upon entry to the page. The entire section is about how one can download forms from the site including a link to Adobe's site for the Acrobat Reader program. Moreover, there is also a link to a search page. On the bottom of this screen shot we see a pencil. As a 53 year old female librarian noted: "The pencil makes you think you're at the end". She was not the only one to have this impression.

In fact, several other users never made it below the pencil, rather, they clicked on the search button and got to the forms only after more clicks. Dozens of users who eventually made it to this page hesitated in confusion and spent as much as 15-20 seconds reading the instructions before scrolling down to notice that they were already on the right page to download the desired tax form.
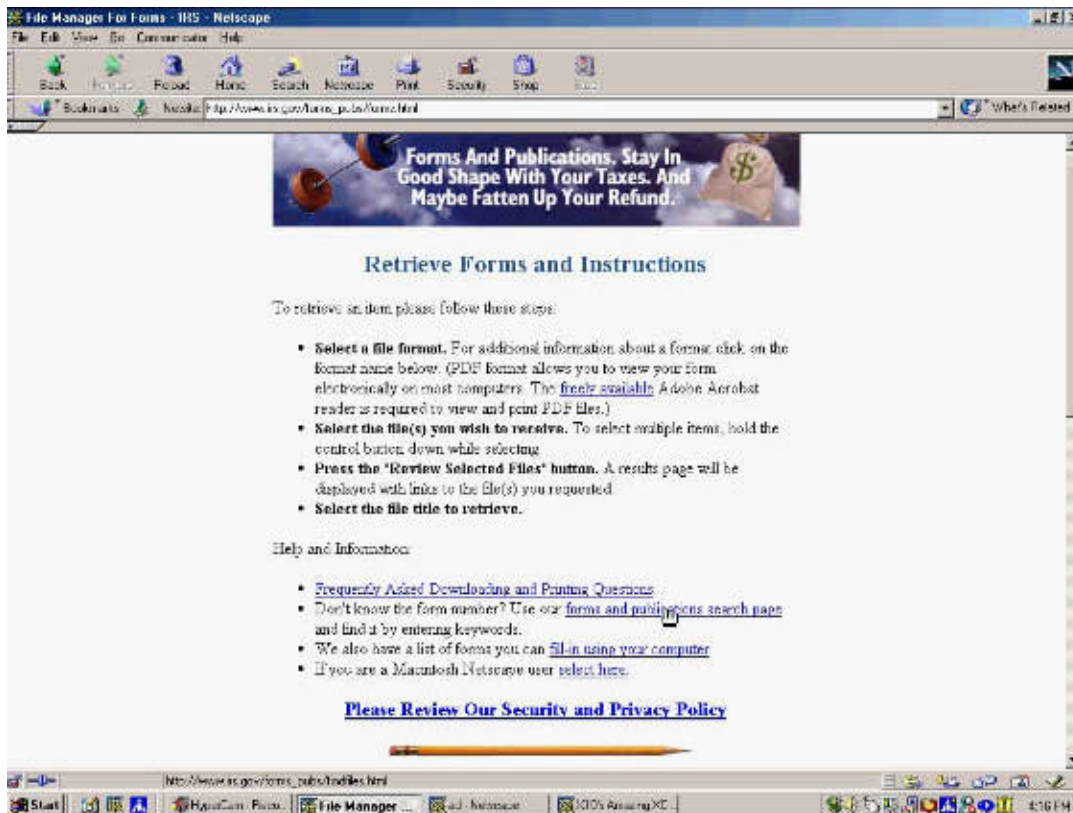

Figure 6. Screenshot of the IRS's Forms page (Fall, 2001)

Ironically, although the IRS has updated its Web site in 2002 and removed the pencil design, the forms section is still not visible "above the fold" on the Forms and Instructions page (see Figure 7 for an example of what the page looks like today). One way to improve this user experience would be to place the forms on top of the page and then have a visible link to Instructions next to it. Putting instructions first – instructions that are in fact irrelevant to many of the users as they already have Adobe Acrobat installed – confuses users because they cannot see that they have reached the target page and wastes their time as they read through irrelevant instructions.
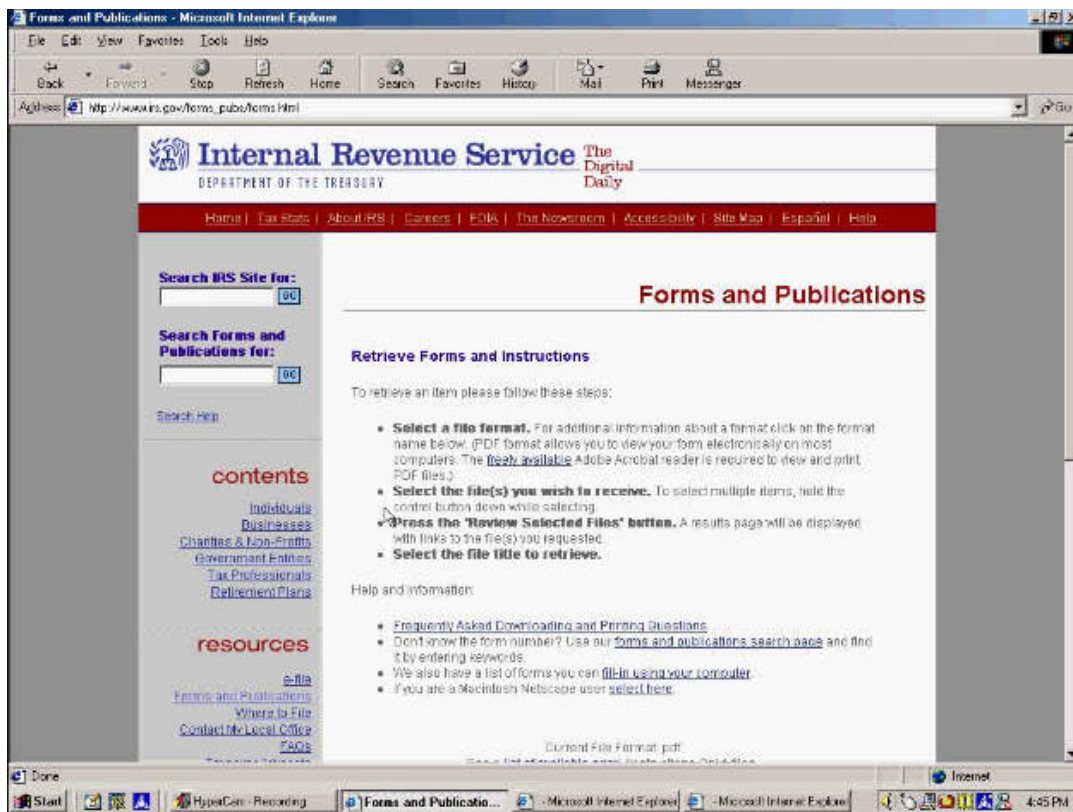
Figure 7. Screenshot of the IRS's Forms and Instructions page (Spring, 2002)

Conclusion

Based on the search patterns of 97 randomly selected Internet users we see that users look for government forms in different ways. For something as widely known as the Internal Revenue Service, a significant number of people will try to access the agency's site directly. Although we may hypothesize that those who know the exact URL of an agency will be quicker to access the desired information, if the site is not organized well and has confusing design elements, directly accessing the site will not make a user more efficient in obtaining the information sought. In fact, there is no statistically significant difference in time-to-completion between those who went to "www.irs.gov" as their first move in this search and those who tried elsewhere probably due to some of the confusion that results from site organization, layout and design.

There are several ways in which the situation can be improved. First, site designers need to be conscious of the effects of their design decisions. In particular, search forms should be available directly on the welcome page of a site. The 2002 version of the IRS's site has a Forms and Publications search button to which several of the respondents from the 2002 portion of the study turned immediately upon arrival

on the site (see Figure 8). Moreover, the focus of a page – in this case the link to the actual forms – should be made clearly visible on the section of the page that will be on the screen of users with modest sized monitors (15" or 17"). Second, a Web site should be set up so if a user types in its address without "www", the user will be redirected to the "www" version of the address. Many sites have implemented this usability feature, why are there sites that do not assist users in this way? Third, users need to be educated about the importance of domain name extensions so they are not derailed by commercial sites that do not necessarily have in their interest to point searchers to the information they seek. Following just these few recommendations would have eliminated the majority of problems users ran into in the case of searching for tax forms in this study, would have shortened the amount of time they spent on the task and would have led to less frustration on their part.

Figure 8. Screenshot of www.irs.gov (Spring, 2002)

References

Benton, Foundation, and Leadership Conference on Civil Rights Education Fund. 2002. "Bringing a Nation Online: The Importance of Federal Leadership." Washington, D.C.: Joint Report.

Bucy, E. 2000. "Social Access to the Internet." *Harvard International Journal of Press/Politics* 5(1):50-61.

Compaine, B. 2001. "Declare the War Won." Pp. 315-335 in *The Digital Divide: Facing a Crisis or Creating a Myth?*, edited by B Compaine. Cambridge, MA: MIT Press.

DiMaggio, Paul, Eszter Hargittai, Russell Neuman, and John Robinson. 2001. "Social implications of the Internet." *Annual Review of Sociology* 27:307-336.

Hargittai, Eszter. In press. "Beyond logs and surveys: In-depth measures of people's Web use skills." *Journal of the American society for information science and technology perspectives*.

Hoffman, DL and TP Novak. 1998. "Bridging the Racial Divide on the Internet." Pp. 390-391 in *Science*. 280.

Howard, P.E.N., L. Rainie, and S. Jones. 2001. "Days and Nights on the Internet: The Impact of a Diffusing Technology." *American Behavioral Scientist* 45(3):383-404.

Jansen, B.J., and U. Pooch. 2001. "A Review of Web Searching Studies and a Framework for Future Research." *Journal of the American Society for Information Science and Technology* 52(3):235-246.

McDonald, Sharon, and Linda Spencer. 2000. "Gender Differences in Web Navigation: Strategies, Efficiency, and Confidence." Pp. 174-181 in *Women, Work, and Computerization: Charting a Course to the Future*, edited by Ellen Balka and Richard. K. Smith. Boston: Kluwer Academic Publishers.

National Telecommunications and Information Administration. 2002. "A Nation Online." Washington, D.C.

Panicware. 2001. "Don't Panic 4.0."

Strover, S. 1999. "Rural Internet Connectivity." Columbia, MO: Rural Policy Research Institute.

Wang, Peiling, William B. Hawk, and Carol Tenopir. 2000. "Users' Interaction with World Wide Web Resources: An Exploratory Study Using a Holistic Approach." *Information Processing and Management* 36(2):229-251.

Wilson, E.J. 2000. "Closing the Digital Divide: An Initial Review: Briefing the President." Washington, D.C.: Internet Policy Institute.