[DigLib] Stanford Digital Library Project
Working Paper

Title: The Proxy is where it's at!
Author: Terry Winograd
Abstract:

Soon, more than 90% of access to the Internet will be through commercial proxies.

The Proxy is where it's at!

I've been thinking about the third-party information uses that we explored with ComMentor, in the light of the current flap about "KidCode." and other schemes for limiting what kids can see on the Net. Many people clearly desire to provide mechanisms for ratings and filters for material on the Internet, seeing these as an alternative to Draconian prohibitions of objectionable material appearing there at all. I was trying to think about how our mechanisms would fit in.

I realized that my own thinking has been dominated by a style of using the Internet that is familiar and prevalent today, but will soon go the way of the build-it-yourself car and the component audio system (i.e., small niche markets). We have been thinking primarily about people controlling access to the Internet from their own PC or workstation. That will not be very common.

* Soon, more than 90% of access to the Internet will be through commercial proxies.

I don't put a date on this, but it's a small number of years. This development will have a serious effect on the design of network information services, of the kind we are developing in the Digital Libraries project. We should think about the implications.

There are two different groups of proxies. One is the firewall proxy now being used by almost all relatively large industrial sites. People using the Internet at work go through their company proxy, and this will become true for smaller companies as well. The other kind of commercial proxy is the "home provider", which individual subscribers use as their access point to the Internet. Today this already applies to the large number of Internet users coming in through on-line services like AOL. There have been some claims that these services will lose their market once people can get onto the net directly. I believe the opposite. If the AOLs and Microsoft Networks of the world do it right, they will increase their market and the ways in which they can make income from new services. They will be the nearly universal way that people connect to the net.

My prediction is that almost everyone who uses the Internet from home will subscribe to a "home provider" (I use the word "home" here as in "home page") as their primary access point to the Internet. This provider will bundle a set of related services, and it is the bundling that gives them the leverage. I will describe the reasons for this in the home market, and then project it back onto the commercial one.

1) Access

In the old on-line world, access meant that you logged in to the provider's computer. But it is moving more towards the point where the physical aspect of the service is primarily dialing in to their modem pool to get on the net, and the computing is done at home. The access aspect will remain necessary for the foreseeable future - we may move from regular phone lines to ISDN, etc. but you need to be able to connect through some kind of centralized hub.

2) 24-hour presence

When you subscribe to a centralized service provider, it can be running its machines for you reliably, connected 24 hours a day to the net. They can be collecting mail, sifting news articles, making your own web pages available to the world, running your shopping agent, etc., etc. For the vast majority of home users, it doesn't make sense to stay connected all the time. If you want a home page on the web, a machine has to be serving it that doesn't disappear when the kids need to make a phone call on the other line. Those of us who have a de facto home provider (the university) don't tend to be aware of how much we take for granted this kind of free service. It isn't there for most people.

3) Filtering and ratings

All the hullabaloo about keeping bad stuff out of kids hands has already been dealt with in a very simple way by services like AOL for their internal materials. They designate which are adult, and they provide the home user with multiple passwords for individual family members and the ability for the account owner to set permissions on any of the passwords. They can do this because all accesses pass through their control mechanism. They could easily do this with the Web as well -- they establish a policy for determining what web sites are good and bad, and when you send a URL through their proxy (which is your only access to the net if you are going through them), they can look at how you are logged in and decide whether to give you the page or not.

This doesn't solve the problem of who does the ratings, but it does make it technically easy to implement the resulting policy. This would be equally true of a service that simply provided a proxy service to anyone on the net who wanted to sign up. The service would keep track of identities, passwords, etc. and you could choose between competing rating schemes as you desire. The proxy service doesn't care what you see, just that it is determined by what you tell them to allow. This whole problem a more serious issue than many in the young-adult academic community take it to be. It will dominate many consumer decisions about whether it is OK to put the household on the net or not. Services that provide a feeling that it is taken care of will dominate those that can't.

4) Payment

In our project we have developed ideas for highly distributed payment systems. But there are definite economies of centralization, both for the payer and payees. Many of the issues are simplified by putting reputable institutions in the middle, who can demand real user information (your name, address, bank account number, etc.) before signing you up, can guarantee payments, carry float, etc. The consumer can get a unified bill, get help with problem transactions, enter package subscriptions that cover multiple services, etc. Providers can make deals for minimal returns, guarantees, percentages of use, etc. that are more flexible when done collectively than with one customer at a time. The whole range of transactions can be managed by a single organizational entity with a large staff of appropriately trained and focused personnel, rather than being scattered in a hodgepodge of organizations around the net.

5) Community interaction

The reason people still use the Well isn't its high tech computer features (which it doesn't have) but the fact that there is a community of people who interact with each other. We can talk about ways to let communities form naturally in a distributed environment, but in fact the successful examples generally have featured someone in a central position who takes responsibility for shepherding the group, often being paid out of proceeds to a general service (such as Well or AOL). There will of course be examples of home-brew groups, but the 90% case will be commercially facilitated ones, which can be most effectively created and managed by a centralized services provider that can maintain a pool of subscribers, collect information about them, keep an eye on the interactions, hire and train staff who understand group facilitation, etc. The kinds of annotation and co-presence mechanisms that we have been talking about in our project all require someone to run and maintain a server that holds and uses the meta-information (sets, groups, people's connections, etc.). This is more easily done by a home service provider than by a standalone enterprise.

6) Advertising and customer information

One big problem with Web advertising today is that you don't know who's reading it or what they're doing. Advertisers want good demographic information on who's reading the ads, and even better, how this correlates with their shopping behavior. This can be used for general evaluation of strategies, or for targeted advertising to those individuals most likely to be interested.

I realize that for many of you, this raises a red flag -- you don't want more advertising in your lives, especially targeted to your consumer habits. You're probably like me -- you throw out all of those glossy color advertising supplements that fill the Sunday newspaper, even before you bring it into the house. But if we were typical, companies wouldn't pay all that money to put those bundles in there. People like advertising, especially if it is engaging and relates to their desires (did you ever look at the Macy's ads on the first few pages of the daily Chronicle?).

By having a proxy provider intermediate between consumers and commercial sites, there are many ways to collect and pass along relevant information. The providers can give advertisers useful demographics, since they have personal information about the users, and also can see what their sequences of actions are, not just the accesses to one site. They could even correlate this with purchase information provided by the companies doing the advertising (who may have records of who bought their products) or acquired from credit card companies. They can collect information from multiple companies and provide users with targeted information ("You might be interested in the sale this week at Fry's...."). They can add advertising on their own as things pass through their proxy, suiting it to the material being sent, the profile of the user, and the budgets of their advertising customers.

They can also provide special deals, discounts, etc. For example, in exchange for getting good consumer demographic information, a company might say that any purchases coming through a home provider will get a 5% discount (or frequent flier miles) - as long as the customer agrees that by doing this they permit the company to collect certain pre-specified information about their profile and reading and buying patterns. Those with a high sensitivity to privacy may find this abhorrent, but for the great majority of people, it is worth giving out non-intimate personal information in exchange for monetary rewards. There needs to be an easy escape (e.g., separate mode or password) for people to use when doing things that might be embarrassing. But if you are dealing with a large home provider, you can have more trust that when you go into "no show mode" they really won't keep records or pass any info about you to the sites you are reading or purchasing from. It is much harder to trust individual companies (are you sure that www.girlsRus.com isn't keeping records of all those accesses?).

7) Updates and new services

Services and applications are changing fast enough that it is a job for any user to keep up with them. This is especially true of the home user who does not have any desire to understand or work with loading software, doing upgrades, etc. They want to have the most current things available, decide whether to pay for those that require more money, and then just have them there. The economics of scale certainly argue that it is easier to upgrade a few dozens, or even hundreds or thousands, of proxy sites, rather than tens of millions of individual and idiosyncratic PC and workstation installations. The home provider will take care of identifying, advertising, installing, maintaining and providing help with all aspects of the software. And this leads us to......

8) Individual service

Actually this last should go first, since it dominates all the others. You have surely heard of the home user who managed to get onto the Internet, ran into something that didn't work as expected, and asked "What number do I call for customer service on the Internet?" It is possible to imagine a world in which all of the different kinds of services mentioned above are offered in a fully distributed way -- you can pick your "ratings provider", your "community connection provider", your "software update service," etc. and mix and match as you go along. This will always appeal to some people -- the same ones who do their own general contracting on their house remodeling, and put together their stereos using components from several manufacturers. But the other 90+% of the market doesn't want to be bothered.

They want to go to one place with all of their questions, complaints, and desires, and to expect someone on the end of the line who doesn't say "Sorry, it doesn't seem to be a browser problem. You'll have to call your ratings provider about that".

* The future of business on the Internet is not a technology business or an information business, but a service business.

The characteristics of service businesses (which favor centralized points of access) will have much more influence on how things are organized than the technical characteristics of the Internet (which favor distributed point-to-point transactions). The social factors will win out over the technical ones.

I am fully convinced that the role of "home service provider" will be a central one, and will cover users in the workplace as well. Company firewall proxies may be managed internally, but the great majority of companies, except for the largest, will buy "wholesale" services from a home provider for many of the above items. This will also include the provider-side of services. Staff and facilities for building web pages, catalogs, and information services will be provided by a large home provider for use by the subscribing company (as EIT is now doing on a smaller scale).

It isn't clear how all this will play out in the marketplace. There are several possibilities:

Natural monopoly
: Whenever there is centralization with positive feedback (the more you have, the more advantageous it is for someone to join), there is a pressure towards the biggest edging out the next-biggest, until there is no next-biggest. This is the dream of the Microsoft Network

Natural oligopoly
: There are reasons having to do with market segmentation and competition that can prevent the feedback loop from settling with a single provider, and lead instead to a small number of very large ones (Ford, GM, and Chrysler; ABC, NBC, CBS and Fox; Windows, MacOS and Unix,...) This is probably the dream of AOL, Compuserve, etc., who might not win in a shoot-out with Bill Gates, but can be among the biggies.

Local providers
: If there is some notion of locality that makes it more effective to focus service on a local set of users, then there can be a relatively large number of providers divided up by this locality. One of the main ideologies of computer networking is that it gets you away from

physical locality - it doesn't matter where you are actually located. Are there other kinds of localities (age groups, professions, cultural backgrounds,....) that would have a strong enough effect on shaping the complex of services that it would be possible to sustain multiple localized services tailored to them?

Mix

Of course these are idealized models. Even in a monopoly or oligopoly there will be special niches for local providers. There will always be some residue of roll-your-own enthusiasts, concentrated in places like Silicon Valley. My prediction is that this will include the university population for a while, but ultimately universities will be handled in the same way as company users, with firewall proxies, services contracted from the big providers, etc.

In doing the Digital Libraries project we are not trying to invent a business plan, but to do research on new technologies. As such we should be guided by what is most interesting technically and most relevant to our testbed. Our testbed is short term and focused on academic and technical users, and as such it will not be strongly affected by these trends. But we also need to keep an eye on the future, so we don't get sidetracked into putting a lot of effort into technical developments that are social or economic dead ends.

--t

Change history:
Original version, both email and HTML, June 24, 1995 by Terry Winograd